

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 01:20:34 ; Search time 1776.4 Seconds  
(without alignments)  
11108.167 Million cell updates/sec

Title: US-09-555-093-1  
Perfect score: 1462  
Sequence: 1 gctcaccacaaatgtgtctgc.....ataaaacctgttttcaa 1462

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	743.4	50.8	773	10	BJ153638
c 2	736.4	50.4	746	10	BJ139067
c 3	736.4	50.4	761	10	BJ129368
c 4	733.8	50.2	763	10	BJ149952
c 5	731.4	50.0	756	10	BJ128883
c 6	727.4	49.8	743	10	BJ128630
c 7	724.8	49.6	742	10	BJ129934
c 8	724.4	49.5	751	10	BJ146824
c 9	724.2	49.5	748	10	BJ134106
c 10	722.2	48.4	750	10	BJ148511
c 11	719.6	48.2	762	10	BJ128407
c 12	707	48.4	731	10	BJ149141
c 13	704.4	48.2	720	9	AU210015
c 14	702.8	48.1	721	10	BJ151779
c 15	701.8	48.0	725	10	BJ152662
c 16	701.8	48.0	727	9	AU213941
c 17	694.6	47.5	719	9	AU213658

c 18	693	47.4	720	10	BJ130441	BJ130441
c 19	682	46.6	705	10	BJ140668	BJ140668
c 20	677.4	46.3	695	10	BJ128927	BJ128927
c 21	675.8	46.2	689	10	BJ131677	BJ131677
c 22	675.2	46.2	686	10	BJ138657	BJ138657
c 23	674.4	46.1	704	10	BJ137106	BJ137106
c 24	665	45.5	689	9	AU209976	AU209976
c 25	662.8	45.3	696	10	BJ122807	BJ122807
c 26	657	44.9	682	10	BJ130694	BJ130694
c 27	641.2	43.9	672	10	BJ151152	BJ151152
c 28	639.2	43.7	671	10	BJ134613	BJ134613
c 29	636.2	43.5	655	10	BJ144819	BJ144819
c 30	635.4	43.5	658	9	AU202072	AU202072
c 31	633.4	43.3	650	9	AU215555	AU215555
c 32	624.4	42.7	627	10	BJ155885	BJ155885
c 33	623.6	42.7	650	10	BJ143820	BJ143820
c 34	621.4	42.5	623	10	BJ120260	BJ120260
c 35	618.6	42.3	639	10	BJ134940	BJ134940
c 36	611.4	41.8	639	10	BJ154428	BJ154428
c 37	600	41.0	637	9	AU218806	AU218806
c 38	598.4	40.9	625	10	BJ136773	BJ136773
c 39	597.4	40.9	614	10	BJ104065	BJ104065
c 40	596.4	40.8	605	9	AU217864	AU217864
c 41	595.8	40.8	618	10	BJ114675	BJ114675
c 42	593.2	40.6	627	10	BJ147665	BJ147665
c 43	592.4	40.5	594	10	BJ103098	BJ103098
c 44	591.6	40.5	618	10	BJ106921	BJ106921
c 45	588.4	40.2	605	10	BJ101261	BJ101261

ALIGNMENTS

RESULT 1  
BJ153638/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
BASE COUNT  
ORIGIN  
Query Match

773 bp mRNA linear EST 24-JAN-2002  
BJ153638 unpublished oligo-capped cDNA library, C. elegans LI stage;  
Caenorhabditis elegans cDNA clone yk1321f12 3', mRNA sequence.  
BJ153638  
EST.  
BJ153638.1 GI:18321623  
Caenorhabditis elegans.  
Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 773)  
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
A complementary view of the C.elegans genome  
Unpublished (2002)  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
1..773  
/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1321f12"  
/clone\_lib="unpublished oligo-capped cDNA library, C.  
elegans LI stage"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"

255 a 136 c 160 g 222 t  
50.8%; Score 743.4; DB 10; Length 773;

Best Local Similarity 99.5%; Pred. NO. 3.3e-127;  
Matches 767; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

```
Qy 690 tcgcatttattccagagatttgcagagataagccagcttggaaagcaattctca 749
Db 773 TCGCATTTATTCAGAGATTGTGCAAGTATGAGCCAGCTTTGAAAGCAATTCICA 714
Qy 750 agattgtaccatacaacatctctatttcacgcgaatgcttccaatgctccgtttctcat 809
Db 713 AGATGTACCATATCAACATCTATTTCACCGCAATGCTTCCATGCTCCGTTCTCAT 654
Qy 810 ggaactgggtcagtcagttcaatttcgaagagatacaaatgaggtacaggtctatc 869
Db 653 GGACGTGTCAGTCAGTTCAATGGGTATTCAAAGAGAAATCAAAATGAGTACAAAGGTCTATC 594
Qy 870 aaagaaatgcatcttcggagcaagaacaattgttgacattggcctgggtattctatc 929
Db 593 AAAGAAATGCAATCTGCGGAGCAAGCAACAATTTGTTGGACATTTGGGCTTGGGTATTCATTC 534
Qy 930 aattgtcttattacaaatgcccacttcgggttgcttatttcatttccacaatgg 989
Db 533 AATTGTTCTTATACCAACATGCCACATTCGGGTGCTTATTTCATATTTCACAATGG 474
Qy 990 gaggaggcctttgtattgctacgtagtcgaacttcaacacataactctgttgataagtc 1049
Db 473 GAGGAGGCGCTTTGATTGCTACGTAGTCACATTTCAACCATTAACCTCTGTTGATAAGTATC 414
Qy 1050 cagcaaatctcgaattttaacaacttcgcgcctctcaaatgttgaccacacgaaca 1109
Db 413 CAGCAATTTCTGCAATTTTAAACAACCTTCGCGGCTCTTCAAAATTTTGACCACACGCAACA 354
Qy 1110 tgactccatctccattcattgattggcttgggtggactcaattatcagacgagcacc 1169
Db 353 TGACTCCATCTCCATTCATTGATTGGCTTTGGGTTGGACTCAATATCAGATCGAGCACC 294
Qy 1170 actgttcccaacaatgcacagtgcgaactcgaatgcttgctggaatattgtgaaagaat 1229
Db 293 ACTGTGTTCCCAACAATGCGACGTTGCAATCTGAAATGCTTGGCATGAAATATGTGAAAGAAAT 234
Qy 1230 ggtgcaagagaaataatctcttccttcctgcgtgcagactacttgcaggtatgcaatga 1289
Db 233 GGTGCAAGAGAAATATCTTCTTACCCTGCTGATGACTACTTTGACGGATATGCAATGA 174
Qy 1290 atttgcaacaattgaaaaatattggtgagcacattcaagctaaagctgcctaaacaattc 1349
Db 173 ATTTGCAACAATTGAAAAATATGGCTGAGCACATTCAGCTAAAGCTGCCTAAACAATCT 114
Qy 1350 ggggttcaaaaagtgttttcttg--tttttaaatattaattcttggaaattattgttt 1407
Db 113 GGGTGTTCAAAAAGTTTTTCTGTGTTTTTTTAAATTTAAATTTCTTTGAAATTAATTTGTTT 54
Qy 1408 tcgctcattcttcctccattccttctggtagataaaacacctgtttt 1458
Db 53 FCCGTCATCTCTCCATCCCTTTTC--GGTAGAAATAAAACCTGTGTTT 4
```

## RESULT 2

BJ139067/c 746 bp mRNA linear EST 23-JAN-2002  
LOCUS  
DEFINITION BJ139067 unpublished oligo-capped cDNA library, C. elegans L1 stage  
Caenorhabditis elegans cDNA clone yK1142e08 3', mRNA sequence.  
ACCESSION BJ139067  
VERSION BJ139067.1 GI:18299233  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans.  
ORGANISM Caenorhabditis elegans.  
Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditiodea  
; Rhabditiidae; Pelodierinae; Caenorhabditis.  
REFERENCE 1 (bases 1 to 746)  
AUTHORS Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.  
and Sugano, S.  
TITLE A complementary view of the C. elegans genome  
JOURNAL Unpublished (2002)

## COMMENT

Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

## FEATURES

source  
1..746  
/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yK1142e08"  
/clone\_lib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"

BASE COUNT 233 a 139 c 156 g 218 t  
ORIGIN

Query Match 50.4%; Score 736.4; DB 10; Length 746;

Best Local Similarity 99.9%; Pred. NO. 6.5e-126;

Matches 737; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 642 ccacaaatgtaattgatcatgacgtgatctgacttggcaccacttttcgatttacc 701
Db 746 CCACAAATGTAATGATCATGACGGTGATATGCACTTGGCACCACATTTTCGCAATTATTC 587
Qy 702 caggagattgtgcaagtataaggccagctttgaaaaagcaattctcaagattgtaccat 761
Db 686 CAGGAGATTGTGCAAGTATAAGGCAGCTTTGAAAAAGCAATTTCTCAAGATTGTACCAT 627
Qy 762 atcaacatctctatttcacccgaatgcttccaatctccgtttctcatgagcgttcagt 821
Db 636 ATCAACATCTCTATTTCACCGCAATGCTTCCAATGCTCCGTTTCATGGAGCTGTCAGT 567
Qy 822 cagttcaatgggtattcacaagagaatcaaatggagtcacaaagctctatcaaaagaaatgc 881
Db 566 CAGTTCAATGGGTATTCAAGAGAGATCAATGAGCTACAAAGTCTATCAAAAGAAATGCAT 507
Qy 882 tctgggagcaagcaacaattgttgacattgggcttgggtatctcatcaattgttcatt 941
Db 506 TCTGGGAGCAAGCAACAATTTGTCGACATTTGGCTTGGGTATTCATCAATTTGTTCTTAT 447
Qy 942 taccacaatgcccacttcgggttgccttatttcattatttcacaaatggagagagccttt 1001
Db 446 TACCAACATGGCCACTTCGGGTGCTTATTTCATTTTCAAAAATGGAGGAGGCCCTTT 387
Qy 1002 tgattgctcacgtagtcactttcaaccataaactctgttgataagtatccagccaattctc 1061
Db 386 TGATTGCTCAGCTAGTCACATTTCAACCATAACTCTGTGTATAAGTATCCAGCAATTCCTC 327
Qy 1062 gaattttaacaaacttcgcctcttcaaatatttgaccacacgcgaacatgactccaatc 1121
Db 326 GAAATTTAAACAACCTTCGCGCTCTTCAAAATTTTGACCACACGCAACATGACTCCATCTC 267
Qy 1122 cattcatgattggcttgggtggagccaattatcagatcgagcaccactgttcccaa 1181
Db 266 CATTCATTGATTGGCTTTGGGGTGAGCTCAATATTCATATCGAGCACCACCTTTGTTCCCAA 207
Qy 1182 caatgccacgttgcgaatcgttcgctgaaatatgtgaaagatggtgcaagaga 1241
Db 206 CAATGCCAGTTGCAATCTGAATGCTTGCATGAATATGTGAAAGAAATGGTCAAGAGA 147
Qy 1242 ataacttcttaccctcgtcgtgactacttgcaggtatgcaatgcaattgtgcaacaat 1301
Db 146 ATAATCTTCTTACCTCGTCGATGACTACTTTGACGGATATGCAATGAAATTTGCAACAAAT 87
Qy 1302 tgaaaaaatatggctagcacattcaagctaaagctgcctaaacaactctgggtgttcaaaa 1361
Db 86 TGAATAATATGGCTGAGCACATTTCAAGCTTAAAGCTTAAAGCTTAAACAATCTGGGGTGTCAAAA 27
```

```

QY 1362 agtttttttctgttttt 1379
Db 26 AGTTTTTCTTGTTTTT 9

RESULT 3
BJ129368/c
LOCUS
DEFINITION
Caenorhabditis elegans cDNA clone yk1029h10 3', mRNA sequence.
ACCESSION
BJ129368
VERSION
BJ129368.1
KEYWORDS
EST.
ORGANISM
Caenorhabditis elegans.
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 761)
AUTHORS
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE
A complementary view of the C.elegans genome
JOURNAL
Unpublished (2002)
COMMENT
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..761
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1029h10"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"

BASE COUNT 248 a 135 c 157 g 220 t 1 others
ORIGIN

Query Match 50.4%; Score 736.4; DB 10; Length 761;
Best Local Similarity 98.8%; Pred. No. 6.5e-126;
Matches 752; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 690 tcgcatttattccaggagatttgcgaagtataaggccagctttgaaaaagcaattctca 749
Db 761 TCGCATTTATTCAGGAGATTGTGCAAGTATAAGGCCAGCTTTGAAAAAGCAATTCTCA 702

QY 750 agattgtaccatatcaacatctctatttcacgcgaatgcttccaatgctccgtttctcat 809
Db 701 AGATTGTACCATATCAACATCTCTATTTCACGCCAATGCTTCCAATGCTCCGTTTCTCAT 642

QY 810 ggactgggtcagtcagttcaatgggtattcaagagaatacaaatggagtagcaaggtctatc 869
Db 641 GGACTGGTCAGTCAGTTCAATGGGTATTCAAGAGAAATCAATGGAGTACAGGTCATC 582

QY 870 aaagaaatgcattctctggagcagcaacaatttggacattgggcttggtattctatc 929
Db 581 AAAGAAATGCATTCTGGGAGCAAGCAACAATTTGTTGGACATTGGGCTTGGGTATTCTATC 522

QY 930 aattgttcttattaccaatgcgaccttcgggttgcttatttcattatttcacaaatgg 989
Db 521 AATTGTCTTATTACACATGGCCACTTCGGGTGCTTATTTCATTATTTCACAAATGG 462

QY 990 gaggggcttttgattgcacgtagtcactttcaaccataactctgttgataagtc 1049
Db 461 GAGGAGGGCTTTTGATTGCTACGAGTACACTTCAACCATCACTCTGTTGATAGTATC 402

```

---

```

QY 1050 cagccaattctogaattttaacaacttcgcgcgtcttcaaattttgacacacgcaaca 1109
Db 401 CAGCCAATTCTCGAATTTTAAACAACCTTCGCCGCTCTTCAAAATTTTGACCACACGCAACA 342

QY 1110 tgactccatctccattcattggttgcttgggtgagctcaattatcagatcgacacc 1169
Db 341 TGACTCCATCTCCATTCAATTGATGTGGCTTTGGGGTGGACTCAATTATCATGATCGACACC 282

QY 1170 acttggtcccaacaatgcacgcttgcaactgaatcgtctgcgtgaaatattgtgaaagaat 1229
Db 281 ACTTGTTCCCAACAATGCCACGTTGCAATCTGAATGCTTGCATGAATATGTGAAGAAT 222

QY 1230 ggtgcaagagagaataatcttccttacctctgcgtgactactttgacgatatgcgaatga 1289
Db 221 GGTGCAAGAGAGAATAATCTTCCTTACCTGCTCGATGACTACTTTTGACGGATATGCAATGA 162

QY 1290 atttgcaacaattgaaaaatagctgcgcacacattcaagctaaagctgcttaacaatct 1349
Db 161 ATTGCAACAATTGAAAAATATGGCTGAGCACATTCAAGCTAAAGCTGCCTAACAATCT 102

QY 1350 ggggtgtcaaaaagttttttcttg--tttttaaatatttaattctttgaaattattgttt 1407
Db 101 GGGTGTTCAAAAAGTTTTTCTTGTTTTTTTAAATTAATCTCTTGAAATATTATTGTTT 42

QY 1408 tcgcattcttctccattcccttcttcgttgtagaataaaa 1448
Db 41 TCGGTATCTTCTCTCCATTCCTCTTTTCGCTAGAAAAATAAA 1

RESULT 4
BJ149952/c
LOCUS
DEFINITION
BJ149952 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone yk1280d08 3', mRNA sequence.
ACCESSION
BJ149952
VERSION
BJ149952.1
KEYWORDS
EST.
ORGANISM
Caenorhabditis elegans.
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 763)
AUTHORS
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE
A complementary view of the C.elegans genome
JOURNAL
Unpublished (2002)
COMMENT
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..763
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1280d08"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"

BASE COUNT 249 a 135 c 159 g 220 t
ORIGIN

Query Match 50.2%; Score 733.8; DB 10; Length 763;
Best Local Similarity 99.2%; Pred. No. 2e-125;
Matches 759; Conservative 0; Mismatches 2; Indels 4; Gaps 2;

QY 690 tcgcatttattccaggagatttgcgaagtataaggccagctttgaaaaagcaattctca 749

```

```
|||||
Db 763 TCGCATTTATTCAGGAGATTGTGCAAGTATPAAGCCAGCTTTGAAAAAGCAATCTCA 704
Qy 750 agatgtaccatatacaatctctatttccacgcgaatcttccaatgctcgtttctcat 809
Db 703 AGATTGTACCATATCAACATCTCTATTTCACGCAATGCTTCCAATGCTCGGTTCTCAT 644
Qy 810 ggcctgggtcagtcagttcaatgggtattcaagagagaatacaaatggatgacaggtctatc 869
Db 643 GGACTGGTCAGTCAGTTCAATGGGTATTCAAGAGAGAAATCAAGTACAGGTCTATC 584
Qy 870 aaagaaatgcattctcgggagcaagcaacaattgttggacattgggcttgggtattctatc 929
Db 583 AAAGAAATGCATCTCGGAGCAAGCAACAATTTGTTGGACATTTGGGCTTGGGTATTCTATC 524
Qy 930 aattgtcttattaccaacatggccacitcgggttgccttatttcaatttccacaaatgg 989
Db 523 AATTGTCTTTATTACCAATGGCCACATTCGGGTTGCTTTATTTCATTTATTCACAAATGG 464
Qy 990 gaggaggcctttgtattgctacgttagtcactttcaacataactctgttgataagtatc 1049
Db 463 GAGGAGGCTTTTGATGCTCAGCTAGTCACCTTTCAACCATAACTCTGTTGATAAGTATC 404
Qy 1050 cagccaattctgaattttaaacaaacttcgcgcctcttcaaatlttgaccacacgcaaca 1109
Db 403 CAGCCAAATCTCGAATTTTAAACAACCTTCGCCGCTCTTCAAAATTTTGACCACGCAACA 344
Qy 1110 tgaactcattcctcattgattgcttgggttgggtggaactcaattatcagatcgagcacc 1169
Db 343 TGACTCCATCTCCATCATTTGATGCTGCTGGGCTTGGGTTGGGACTCAATTTATCAGATCGAGCACC 284
Qy 1170 actgttcccaacaatgcacgctgcaatctgaatgcttgcgtggaatattgtgaagaat 1229
Db 283 ACTGTTCCCAACAATGCCACGTTGCAATCTGAAATGCTTGCAATGAATATGTGAAAGAAT 224
Qy 1230 ggtcgaagagaataatcttctacctcgtcgtgactactcttggacggatagcaatga 1289
Db 223 GGTGCAAGAGAAATATCTCTTACCTCGTGGATGACTTTTGACGGATATGCAATGA 164
Qy 1290 atttgcacaattgaaaaatattgctgagcacattcaagctaaagctgcctaaacaatct 1349
Db 163 ATTTGCACAAATGAAAAATATGGCTGAGCACATTCAGACTAAAGCTGCCCTAAACAATCT 104
Qy 1350 ggggttccaaaagtttttcttg--tttttaaatatttaattcttggaaattattgttt 1407
Db 103 GGGTGTTCAAAAAGTTTTTCTTGTGTTTTTTTAAATTTAAATCTTTTGAAATTTATTGTTT 44
Qy 1408 tccgcatcttccctccattcccttttctgttagaataaaacct 1452
Db 43 TCCGTCATCTTCTCCCATTCCTTTT--CGGTAGAAATTAACACCT 1

RESULT 5
BJ12883/C
LOCUS BJ128883 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1024a12 3', mRNA sequence.
ACCESSION BJ128883
VERSION BJ128883.1 GI:18289040
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 756)
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
A complementary view of the C.elegans genome
Unpublished (2002)
Contact: Tadasi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
```

```
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .756
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1024a12"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"

BASE COUNT 248 a 135 c 157 g 216 t
ORIGIN

Query Match 50.0%; Score 731.4; DB 10; Length 756;
Best Local Similarity 99.6%; Pred. No. 5.4e-125;
Matches 744; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 689 ttcgcatttattccaggagattgtgcaagtataaggccagctttgaaaaagcaattctc 748
Db 756 TTCGCATTTATTCAGGAGATTGTGCAAGTATPAAGCCAGCTTTGAAAAAGCAATCTCT 697
Qy 749 aagattgtaccatatacaatctctatttccacgcgaatgcttccaaatgctcgtttctca 808
Db 696 AAGATTGTACCATATCAACATCTCTATTTCACGCAATGCTTCCAATGCTCGGTTCTCA 637
Qy 809 tggactgtcagtcagtttcaatgggtatttcaagagagaatacaatggagtaagaagctctat 868
Db 636 TGGACTGGTCAGTCAGTTCAATGGGTATTCAAGAGAGAAATCAAAATGAGTACAAAGCTAT 577
Qy 869 caagaaatgattctgggagcaagcaacaattgttgacattgggttgggtattctat 928
Db 576 CAAAGAAATGCAATCTTGGGAGCAAGCAACAATTTGTTGGACATTTGGGCTTGGGTATTCTAT 517
Qy 929 caattgttcttaccacaatcgccacttcgggttgccttatttatttatttccacaatg 988
Db 516 CAATTGTTCTTATTACCAACATGGCCACTTCGGGTGCTTATTTCATTATTTCACAAATG 457
Qy 989 ggaggagccttttgattgctcagtagtcaactttcaaccataaactcgttggataaagtat 1048
Db 456 GGAGGAGGCTTTTGATTGCTCAGCTAGTCACTTTCAACCATAACTCTGTTGATAAGTAT 397
Qy 1049 ccagccaattctcgaattttaaaaaacttcgccgctcttcaaattttgcacacgcgaac 1108
Db 396 CCAGCCAATTCGAAATTTTAAACAACCTTCGCCGCTCTTCAAAATTTTGACCACAGCAAC 337
Qy 1109 atgactccatctccattcattgattggcttgggtggagctcaattatcagatcgagcac 1168
Db 336 ATGACTCCATCTCCATTCATTGATTGGCTTTGGGCTGGACTCAATTTATCAGATCGAGCAC 277
Qy 1169 cacttgttcccaacaatgccagttgcaatctgaaatgcttgcgtgaaataatgtgaaagaa 1228
Db 276 CACTTGTTCCTCCAAACAATGCCAGTTCGAATCTGAATGCTTGCATGAAATATGTGAAAGAA 217
Qy 1229 tgggtcgaagagaataatcttccctacctcgtcgtcagatcactttgacggatatgcaatg 1288
Db 216 TGGTGCAAGAGAAATAATCTTCCCTTACCTCGCTGATGACTACTTTTGACGGGATATGCAATG 157
Qy 1289 aatttgcacaattgaaaaatattgctgagcacattcaagctaaagctgcctaaacaatc 1348
Db 156 AATTGTCAACAATTGAAAAATATGGCTGAGCACATTTCAAGCTAAAGCTGCCTAAACAATC 97
Qy 1349 tgggtgttcaaaaagtttttcttg--tttttaaatatttaattcttggaaattatttgtt 1406
Db 96 TGGGTGTTCAAAAAAGTTTTTCTTGTGTTTTTTTAAATTTAAATCTTCTTGAATTTATTGTT 37
Qy 1407 ttcgcatcttccctccattcccttt 1433
Db 36 TCCGTCATCTTCTCCCATTCCTTTT 10
```



QY	1110	tgactccatctccatctcattgattggctttgggggtggactcaattatccagatcagacacc	1169
DB	323	tgactgccatctccatctcattgattggctttgggggtggactcaattatccagatcagacacc	264
QY	1170	actgtgtcccaacaatgccacgtgccaatctgaaatgcttgcgtgaaatattgtgaaagaat	1229
DB	263	actgtgtgtcccaacaatgccacgttgcgaatctgaaatgcttgcgtgaaatattgtgaaagaat	204
QY	1230	ggtgcaaaagagaataatcttccttacctcgtcgtgactgactctttgacggatatgcaatga	1289
DB	203	ggtgcaaaagagaataatcttccttacctcgtcgtgactgactctttgacggatatgcaatga	144
QY	1290	atttgcaacaattgaaaaatatgctgagcacattcaagctaaagctccctaaacaattct	1349
DB	143	atttgcaacaattgaaaaatatgctgagcacattcaagctaaagctccctaaacaattct	84
QY	1350	gggtgttccaaaaagtgttttcttctg--tttttaaaatttaattctttgaaattattttgttt	1407
DB	83	gggtgttccaaaaagtgttttcttctggttttttttaaaatttaattctttgaaattattttgttt	24
QY	1408	tcggtccattcttcctccattccc	1430
DB	23	tccggtcattcttcctccattccc	1
RESULT	7		
LOCUS	BJ129934/c		
DEFINITION	BJ129934	742 bp	mRNA linear EST 23-JAN-2002
ACCESSION	BJ129934	unpublished oligo-capped cDNA library, <i>C. elegans</i> Ll stage	
VERSION	BJ129934.1	GI:18290091	
KEYWORDS	EST.		
SOURCE	Caenorhabditis elegans.		
ORGANISM	Caenorhabditis elegans.		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea		
AUTHORS	Kohara, Y., Shin, I.-T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.		
TITLE	A complementary view of the <i>C. elegans</i> genome		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Tadasu Shin-I Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp. Location/Qualifiers 1. 742 /organism="Caenorhabditis elegans" /strain="N2" /db_xref="taxon:6239" /clone="yk1036d04" /clone_lib="unpublished oligo-capped cDNA library, <i>C. elegans</i> Ll stage" /sex="hermaphrodite" /tissue_type="whole animal" /dev_stage="L1"		
FEATURES	source		
BASE COUNT	242 a	132 c	156 g 212 t
ORIGIN			
Query Match	49.6%	Score 724.8;	DB 10; Length 742;
Best Local Similarity	99.5%;	Pred. No. 8.9e-124;	
Matches	738; Conservative	0; Mismatches	2; Indels 2; Gaps 1;
QY	690	tgcgattttccaggagatttgccaagtataaggccagcttggaaaaagcaattctca	749
DB	742	tgcgattttccaggagatttgccaagtataaggccagcttggaaaaagcaattctca	683
QY	750	agattgtaccatatcaacatctctatttcacgcgaatgtcttccaatgctccgtttctcat	809

RESULT	6	743 bp	linear	EST 23-JAN-2002
BJ128630/c				
LOCUS	743 bp	linear		
DEFINITION	BJ128630	unpublished oligo-capped cDNA library, C. elegans L1 stage		
ACCESSION	BJ128630	unpublished oligo-capped cDNA clone yk102la08 3', mRNA sequence.		
VERSION	BJ128630			
KEYWORDS	EST.			
SOURCE	BJ128630.1	GI:18288787		
ORGANISM	Caenorhabditis elegans.			
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea			
AUTHORS	1 (bases 1 to 743)			
TITLE	Kohara, Y., Shin, I., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.			
COMMENT	and Sugano, S.			
FEATURES	A complementary view of the C.elegans genome			
source	Unpublished (2002)			
	Contact: Tadashi Shin-i			
	National Institute of Genetics			
	1111 Yata, Mishima, Shizuoka 411-8540, Japan			
	Tel: 81-559-81-6856			
	Fax: 81-559-81-6855			
	Email: tshini@genes.nig.ac.jp.			
	Location/Qualifiers			
	1. 743			
	/organism="Caenorhabditis elegans"			
	/strain="N2"			
	/db_xref="taxon:6239"			
	/clone="yk102la08"			
	/clone_lib="unpublished oligo-capped cDNA library, C.			
	elegans L1 stage"			
	/sex="hermaphrodite"			
	/tissue_type="whole animal"			
	/dev_stage="L1"			
BASE COUNT	243 a	132 c	156 g	212 t
ORIGIN				
Query Match	49.8%	Score 727.4;	DB 10;	Length 743;
Best Local Similarity	99.6%	Pred. No. 3e-124;		
Matches	740;	Conservative	0;	Mismatches 1; Indels 2; Gaps 1;
Qy	690	tcgatttattccaggagattgtgcaagtataaggccagcttggaaaaagcaattctca	749	
Db	743	TCGCATTTATTCAGGAGATTTGTGCAAGTATAGGCCAGCTTTGAAAAAGCAATTC	684	
Qy	750	agattgtaccatatacaacatctctatttcacgcgaatgcttccaatgctcgtttctcat	809	
Db	683	AGATTGTACCATATCAACATCTCTATTTCACCGCAATGCTTCCAAATGCTCGTTTCAT	624	
Qy	810	ggactgttcagtcagttcaatgggtatttcaagaagaaatcaaatggatcaaggtctatc	869	
Db	623	GGACTGGTCACTCAGTTCAATGGGTATTTCAAAGAGAATCAATGGAGTACAAGCTATC	564	
Qy	870	aaagaaatgcattctggagcaagcaacaattgttgacattggccttggtattctatc	929	
Db	563	AAAGAAATGCATTTGGGAGCAGCAGCAATGTTGGACATTTGGGCTTGGGTATTCATC	504	
Qy	930	aattgtttattaccacaatggccacttcgggtgtgcttatttatttatttcacaaatgg	989	
Db	503	AATTGTTCTTATACCAACATGGCCACTTCCTGGTGTCTATTTCATTTATTCACAAATGG	444	
Qy	990	gagaggcccttttgattgctcaactagtcacttcaaccataaactctgttgatgaatc	1049	
Db	443	GAGGAGGCGCTTTTGATTGCTCAGCTAGTCACTTTTCAACCAATAACTCTGTTGATAGTATC	384	
Qy	1050	cagccaattctcgaatttttaaacaaacttcgcgccttcctcaattttgacacacgcgaaca	1109	
Db	383	CAGCCAATTCFCGAATTTTAAACAACTTCGCGCGCTCTTCAAAATTTTGACCAACGCAACA	324	

```
Db 682 AGATTGTACCATATCAATCTCTATTTCCAGCGCAATGCTTCCATGCTCGTTTCTCAT 623
Qy 810 gactgggtcagctcaatgggtattcaagagaaatcaaatggagtacaggtctctac 869
Db 622 GCAGTGGTCACTCAATGGGTATTCAAAGAGAATCAATGAGGACAAAGGTCTATC 563
Qy 870 aaagaaatgctattctgggagcaagcaacaattgttgacattgggtctgggtattctac 929
Db 562 AAAGAAATGCATTCGGGAGCAAGCAACAATTTGGGACATTTGGCTTGGGTATTCATC 503
Qy 930 aatgttcttattaccacacatggcgaacttcgggttggtctattctatttccaaatcg 989
Db 502 AATGTTCTTATTACCAACATGGCCACTTCGGGTGGTTCATTTATTTATTCACAAATGG 443
Qy 990 gagagggcttttgattgctcagctagtcacttccaaacataaactctgtgataagtc 1049
Db 442 GAGGAGGCTTTTGTATGCTCAGCTAGTCACTTTCAACCAATAACCTCTGTTGATAGTATC 383
Qy 1050 cagcaaatctcgaattttaacaacttcgcgcgtctctcaaattttgaccacacgcaaca 1109
Db 382 CAGCCAATTCGAAATTTAAACAACATTCGCGCTCTTCAAAATTTTGACCACAGCAACA 323
Qy 1110 tgactccatctccattcattgattggttgggttgactcaaatattcagatcgagcacc 1169
Db 322 TGACTCCATCTCCATTCATTTGATTTGGCTTGGGTGGACTCAATTTATCAGATCGAGCACC 263
Qy 1170 actgttctccaaatgcccgttgcaatctgaatgcttcgctgaaatctgtgaaagaaat 1229
Db 262 ACTGTTCCCAACAATGCCAGCTGGCAATCTGAATGCTTGCAATGATGTAAGAAAT 203
Qy 1230 ggtgcaagagaaataatctctaccctgcgcgtgactgactcttgacggatgataatga 1289
Db 202 GGTGCAAGAGAAATTAATCTCTTACCTCGTCTGATGACTACTTTGACGGATATGCAATGA 143
Qy 1290 atttgcaacaattgaaataatgctgagacacattcaagctaaagctgcctaaacaatct 1349
Db 142 ATTTGCAACAATGAAATAATGCTGAGCACATTCAAAGCTTAAAGCTGCTTAAACAATCT 83
Qy 1350 ggggttcaaaaagtctttctctg--tttttaaaatttaattcttgaattattgttt 1407
Db 82 GGGGTCTCAAAAAGTTTTTCTGTTTTTTTTTAAATTTAAATCTTTGAAATATTGTTTT 23
Qy 1408 tccgtcattctctccattcc 1429
Db 22 TCCGTCATCTCTCCCATTC 1
```

```
RESULT 8
BU146824/c
LOCUS
DEFINITION BU146824 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone yk1239a08 3', mRNA sequence.
ACCESSION BU146824
VERSION BU146824.1 GI:18306990
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
1 (bases 1 to 751)
and Sugano,S.
A complementary view of the C.elegans genome
Unpublished (2002)
Contact: Tadasu Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
```

FEATURES

```
source 1. .751
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1239a08"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
BASE COUNT 243 a 133 c 154 g 218 t 3 others
ORIGIN
Query Match 49.5%; Score 724.4; DB 10; Length 751;
Best Local Similarity 98.5%; Pred. No. 1.1e-123;
Matches 740; Conservative 0; Mismatches 9; Indels 2; Gaps 1;
Qy 700 tccaggagattgtgcaagtataaggccagctttgaaaaagcaattctcagattgtacc 759
Db 751 TCCAGNAGATTGTGCAAGTATAAGGCCAGCTTTGAAAAAGCAATTCTCAAGATTGTACC 692
Qy 760 atatacaatctctatttaccgcaatgcttccaatgctccgtttctcatggaactgttca 819
Db 691 ATATCAACATCTNTATTTACCGCAATGCTTCCAATGCTCCGTTTCTCATGGACTGGTCA 632
Qy 820 gtccagttcaatgggtattcaagagaaatcaaatggagtacaaggtctatcaaaagaaatgc 879
Db 631 GTCAGTTCAATGGGTATTCAAAGAGAAATCAANTGGAGTACAAAGGTCTATCAAGAAATGC 572
Qy 880 attcggagcaagcaacaattgttgacattgggcttgggtattctctatcaattgttctt 939
Db 571 ATTCTGGGAGCAAGCAACAATTTGTGGACATTTGGCTTGGGTATTTCTATCAATTGTTCTT 512
Qy 940 attcacaacatggccacttcgggttctatttatttatttatttatttatttatttattt 999
Db 511 ATTACCAACATGGCCACTTCGGGTGCTTATTTTCATTTTACAAAATGGAGGAGGCT 452
Qy 1000 ttgtatgctcacgtagtcactttcaaccataactctgtgtgataagtaagtaagtaagta 1059
Db 451 TTTGATTTGCTCAGTGTACACTTTCAACCAATACTCTGTTGATTAAGTATCCAGCAATTC 392
Qy 1060 tcgaattttaacaacttcgcgcgtcttcaaattttgaccacacgcaacatgactccatc 1119
Db 391 TCGAATTTTAAACAACATTCGCCGCTCTTCAAAATTTTGACCAACACACATGACTCCATC 332
Qy 1120 tceattcattgattggttgggttgactcaattatcagatcgagcaccactgttctcc 1179
Db 331 TCCATTTCAATTTGGGCTTTGGGTGGACTCAATTTATCAGATCGAGCACCCTTTGTTCCC 272
Qy 1180 acaatgcccgcttgcaatctgaatgcttgcgtgaaatattgaaagaatggtgcaaga 1239
Db 271 AACAAATGCCAGTTGCAATCTGAATGCTTGAATGCTGTAAGTATCCAGCAAGA 212
Qy 1240 gaataattctcttaccctcgtcgatgactacttgaacggtatgataatgaatttgcaca 1299
Db 211 GAATAAATCTTCTTACCTCGTCGATGACTACTTTGACGGATATGCAATGAATTTGCAACA 152
Qy 1300 attgaaaaatatgctgagcaacatcgaagctcaagctcctaaacaatctgggtgttcaa 1359
Db 151 ATTGAAAAATATGCTGAGCACATTCAGCTTAAAGCTGCCCTTAAACAATCTCGGGTGTCAA 92
Qy 1360 aaagtctttctctg--tttttaaatataattctttgaaattattgttttccgtcattc 1417
Db 91 AAAGTTTTTCTTCTTTTTTTTTTAAATTTAAATTTCTTTGAAATTTATTTGTTTCCGTCATTC 32
Qy 1418 ttctccattccctttcttctgtagaataaa 1448
Db 31 TTCTCCATTCCTCTTTCGGGTAGAAAAATAA 1
RESULT 9
BU134106/c
```

LOCUS BJI34106 748 bp mRNA linear EST 23-JAN-2002  
DEFINITION BJI34106 unpublished oligo-capped cDNA library, C. elegans L1 stage  
Caenorhabditis elegans cDNA clone yk1084g01 3', mRNA sequence.  
ACCESSION BJI34106  
VERSION BJI34106.1 GI:18294263  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans.  
ORGANISM Caenorhabditis elegans  
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
AUTHORS 1 (bases 1 to 748)  
TITLE Kohara,Y., Shin-I., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
JOURNAL A complementary view of the C.elegans genome  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
FEATURES  
source  
LOCATION/Qualifiers  
1..748  
/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1084g01"  
/clone\_lib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"  
BASE COUNT 244 a 134 c 156 g 214 t  
ORIGIN  
Query Match 49.5%; Score 724.2; DB 10; Length 748;  
Best Local Similarity 99.3%; Pred. No. 1.1e-123;  
Matches 738; Conservative 0; Mismatches 3; Indels 2; Gaps 1;  
Qy 697 tattccaggagattgtgcagataagtcagcagcttggaaagcaattctcaagattgt 756  
Db 748 TATTCAGGAGATTTGTCAAGTATATAAGCCAGCTTTGAAAAGCAATTCAGATTGT 689  
Qy 757 accatatcaacatctctattcacgcgaatgtctccaatgtcttccattctcatgactgg 816  
Db 688 ACCATATCAACATCTCTATTACCGCAATGCTTCCAAATGCTCCGTTTCTCATGGACTGG 629  
Qy 817 tcagtcagttcaatgggtattcaagagaatacaaatggagtaaggtctatcaagaaa 876  
Db 628 TCAGTCAGTTCAATGGGTATTCAAGAGCAATCAATGAGTACAGGCTATCAAGAAA 569  
Qy 877 tgcattctggagagaagcaacaattgttggacattgggctgggtattctatcaattgtt 936  
Db 568 TGCATTCTGGGAGCAAGCAACAATTTGTGACATTTGGGCTTTGGGTATTCTATCAATTGTT 509  
Qy 937 cttattaccaacatggccacttcgggtgtctatttcaattatttcaaaaatggagagg 996  
Db 508 CTTATACCAACATGGCCACTTCGGGTTGCTATTATTCATTATTCACAAATGGGAGGAGG 449  
Qy 997 ccttttgattgtcacgtagtcactttcaaccataactctgttgataagatccagccaa 1056  
Db 448 CCTTTGATTGCTCAGGTAGTCACTTTCAACCATAACTCTGTTGATAAGTATCCAGCCAA 389  
Qy 1057 ttctogaatttaacaacttcgcgcgtcttcaaatatttgaccacacacacacactgcc 1116  
Db 388 TTCTCGAATTTAAACAACCTTCGGCGCTCTTCAAAATTTTGACCACACGCAACATGACTCC 329  
Qy 1117 atctccattcattgattggtttgggtggactcaattatcagatcgagaccacttgtt 1176  
Db 328 ATCTCCATTTCATTGATGGCTTTGGGGTGGACTCAATTTATCAGATCGAGCCACTTGT 269

Qy 1177 cccacaatgccacgttgaacttgaatgcttgcgtgaaatgtgaaagaatgggtgcaa 1236  
Db 268 CCCACAAATGCCACGCTGCAATCTGAATGCTTGCATGAATATGTCAAGAATGGTGCAA 209  
Qy 1237 agagaataattcttcttacctcgtcgtgactacttcttgacgagatgcaaatgaattgca 1296  
Db 208 AGAGAATAATCTTCCTTACCTCGTGCATGACTACTTTGACGGATATGCAATGAATTGCA 149  
Qy 1297 acaattgaaaaatatggctgagcacatcaactcaagctcaagctgctcaacaatctgggtgt 1356  
Db 148 ACAATTGAAAATATGGCTGAGCACATTTCAAGCTAAGCTGCTTAACAAATCTGGGTGTT 89  
Qy 1357 caaaaagtgttttcttg--tttttaaaatttaattcttcttgaattattgttttccgtca 1414  
Db 88 CAAAAGTTTTTCTGTTTTTTTAAATTTAATCTTTGAAATTTATTTGTTTTCGTGCA 29  
Qy 1415 ttctctccattccctttctg 1437  
Db 28 TTCTTCTCCATTCGCTTTTCGG 6  
RESULT 10  
BJI48511/c  
LOCUS BJI48511 750 bp mRNA linear EST 24-JAN-2002  
DEFINITION BJI48511 unpublished oligo-capped cDNA library, C. elegans L1 stage  
Caenorhabditis elegans cDNA clone yk1263h03 3', mRNA sequence.  
ACCESSION BJI48511  
VERSION BJI48511.1 GI:18316496  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans.  
ORGANISM Caenorhabditis elegans  
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
AUTHORS 1 (bases 1 to 750)  
TITLE Kohara,Y., Shin-I., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
JOURNAL A complementary view of the C.elegans genome  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
FEATURES  
source  
LOCATION/Qualifiers  
1..750  
/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1263h03"  
/clone\_lib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"  
BASE COUNT 244 a 134 c 156 g 216 t  
ORIGIN  
Query Match 49.4%; Score 722.2; DB 10; Length 750;  
Best Local Similarity 99.3%; Pred. No. 2.7e-123;  
Matches 736; Conservative 0; Mismatches 3; Indels 2; Gaps 1;  
Qy 699 ttccaggagattgtgcagataagtcagcagcttggaaagcaattctcaagattgtac 758  
Db 750 TTCCAGGAGATTTGTGCAAGTATAAGCCAGCTTTGAAAAGCAATTCCTCAAGATTGTAC 691  
Qy 759 catatcaacatctctatttcaccgcaatgcttccaatgctcgttctcatgactggtc 818  
Db 690 CATATCAACATCTCTATTTCACCGCAATGCTTCCAAATGCTCCGTTTCTCATGACTGGTC 631  
Qy 819 agtcggttcaatgggtatttcaagagaaatcaaatggagtacaagggtctatcaagaatg 878

Db 630 AGTCAGTTCAATGGGTATTCAAGAGAAATCAATGGAGTACAAGGCTCATCAAGAAATG 571  
Qy 879 cattctggagcaagcaaatgttgacattggcttggttgcattctatcaattgttct 938  
Db 570 CATTCGGGAGCAACCAATTTGGACATTTGGCTTGGGTATTCATCAATGTCT 511  
Qy 939 tattaccaacatggccacttgcgttgcttatttcatttccaaatggagagagcc 998  
Db 510 TATTTACCAACATGGCCACTTGGCTGGTGTCTTATTTCAATATTCACAAATGGAGAGGCC 451  
Qy 999 ttctgattgcacgtatcactttcaaccataactctgttgataagtatccagccaatt 1058  
Db 450 TTTTGATTGCTCAGCTAGTCACTTTCAACCATAACTCGGTTGATAAGTATCCAGCCAAAT 391  
Qy 1059 ctcgaaatttaaaacacttcgcgtctcttcaaatttgaccacacgcacactgactccat 1118  
Db 390 CTCGAATTTTAAACAACCTTCGCCGCTCTTCAAAATTTGACCACACGCAACATGACTCCAT 331  
Qy 1119 ctccattcattgattggcttgggtggactcaattatcagatcgagcaccacttgttcc 1178  
Db 330 CTCCAATTCATTGATGGCTTTGGGTGGACTCAATTTATCAGATCGAGCACCACCTTGTTC 271  
Qy 1179 caacaatcccagcttgaatcgttgcgtgaaatatgtgaaagaaatgggtgcaag 1238  
Db 270 CAACAATGCCAGCTTGGCAATTCGAATGCTTCATGAAATATGTGAAGAAATGGTGCAAG 211  
Qy 1239 agaatcaatctctcttaccctgcgatgactactttgacggatgcaatgaattgtgaac 1298  
Db 210 AGAATAATCTCCCTACCTCGTGCATGACTACTTTGACGGATGCAATGAATTTGCAC 151  
Qy 1299 aattgaaaaatatggctgagcacattcaagctaaagctgcctaaacaaatctgggtgttca 1358  
Db 150 AATTGAAAAATATGGCTGAGCACATTCAGCTAAAGCTGCCTAAACAATCTGGGTGTCA 91  
Qy 1359 aaagttttttcttg--tttttaaatatttcttgaattatttcttcttcgctcatt 1416  
Db 90 AAAGTGTTCCTGTTTTTAAATTAATCTTTGAAATATTATTTGTTTCGCTCAAT 31  
Qy 1417 ctctccctcattccctttctg 1437  
Db 30 CTTCCTCCATTCCTTTTCGG 10

RESULT 11  
BJ128407/c  
LOCUS  
DEFINITION BJ128407 762 bp mRNA linear EST 23-JAN-2002  
Caenorhabditis elegans cDNA clone yk1018d12 3', mRNA sequence.  
ACCESSION BJ128407  
VERSION 1  
KEYWORDS EST  
SOURCE BJ128407.1 GI:18288564  
ORGANISM Caenorhabditis elegans.  
REFERENCE 1 Rhabditidae; Peleoderinae; Caenorhabditis.  
AUTHORS Kohara,Y., Shin-I., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
TITLE A complementary view of the C.elegans genome  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshin@genes.nig.ac.jp.  
Location/Qualifiers  
1. .762  
/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"

BASE COUNT 250 a 136 c 157 g 219 t  
ORIGIN  
Query Match 49 2% Score 719.6; DB 10; Length 762;  
Best Local Similarity 99.0%; Pred. No. 8e-123;  
Matches 756; Conservative 0; Mismatches 4; Indels 4; Gaps 3;  
Qy 690 tcgcatatttcccaggagattgtgcaagtataagccagcttggaaaagcaattctca 749  
Db 762 TCGCATTTATTCAGGAGATTTGTGCAAGTATAAGCCAGCTTTGAAAGCAATTCTCA 703  
Qy 750 agattgtaccatacaacatctctatttcacgcgaatgtctccaatgctccgtttctcat 809  
Db 702 AGATTTTACCATATCAACATCTCTATTTACCGCAATGCTTCCAATGCTCCGTTCTCAT 643  
Qy 810 gactgtgcactcagttcaatgggtatttcaagagaaatcaaatggagtacaagctctac 869  
Db 642 GGACTGGTCACTCAGTTCAATGGGTATTTCAAGAGAAATCAAAATGGAGTACAGGTCTATC 583  
Qy 870 aaagaaatgcattctggagcaagcaacaattgttgacattgggcttgggtattctatc 929  
Db 582 AAAGAAATGCATTTCTGGGAGCAAGCAACAATTTGGACATTTGGCTTGGGTATTCATC 523  
Qy 930 aattgttcttattaccacaatggccacttcgggttgcgttatttctatttccaaatgg 989  
Db 522 AATTGCTTCTTATACCAACATGGCCACTTCGGGTGCTTATTTCAATTTATTCACAAATGG 463  
Qy 990 gaggagccttttgaatgctcacttagctcacttcaaccataactctgttgaatgattc 1049  
Db 462 GAGGAGGCGCTTTTGATGCTCAGGTAGTCACTTCAACCAATCTCTTTGATAGTATC 403  
Qy 1050 cagcaaatcttcgaaattttaacaacttcgcgcctcttcaaattttgaccacacgcaaca 1109  
Db 402 CAGCCAATTCGGAATTTTAAACAA-TTCGCGCGCTCTTCAATTTTGACCACACGCAACA 344  
Qy 1110 tgactccatctccattcattgattggcttgggtggactcaatatacagatcagagacc 1169  
Db 343 TGACTCCATCTCCATTCATTTGCTTGGCTTGGGTGGACTCAATTTATTCAGATCGAGCAC 284  
Qy 1170 actgtttcccaaatgccacgttgcaatcgtgactgctgctgcaaatatgtgaaagaat 1229  
Db 283 ACTGTGTTCCCAACAAATGCCAGTTCGCAATCTGAATGCTTGCATGAAATATGTGAAAGAA 224  
Qy 1230 ggtgaaagagaataactctcttaccctcgtcgtgactactttgacggatatgcaatga 1289  
Db 223 GGTGCAAGAGAAATTAATCTCTTACCTCGTCGATGACTACTTTGACGGATATGCAATGA 164  
Qy 1290 atttgcaacaattgaaaaaataggctgagcacattcaagctaaagctgcctcaacaatct 1349  
Db 163 ATTTGCAACAATTTGAAAAATATGGCTGAGCACATTTCAAGCTTAAAGCTGCCCTTAAACAATCT 104  
Qy 1350 ggggttccaaaagtttttctcttg--tttttaaatatttcttgaattatttctt 1407  
Db 103 GGGTGTTCAAAAGAGTTTTTCTGTTCTTTTTTAAATTTAAATCTTTGAAATATTATTTGTTT 44  
Qy 1408 tccgtcattctctccattccctttctctgtgtagaataataaac 1451  
Db 43 TCCGTCATCTTCTCCATTCCTTTTC-GGTAGAAATAAAGC 1  
RESULT 12  
BJ149141/c  
LOCUS  
DEFINITION BJ149141 unpublished oligo-capped cDNA library, C. elegans L1 stage  
Caenorhabditis elegans cDNA clone yk1271c01 3', mRNA sequence.  
ACCESSION BJ149141

731 bp mRNA linear EST 24-JAN-2002  
BJ149141  
BJ149141 unpublished oligo-capped cDNA library, C. elegans L1 stage  
Caenorhabditis elegans cDNA clone yk1271c01 3', mRNA sequence.  
ACCESSION BJ149141

VERSION BJI49141.1 GI:18317126  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans.  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
REFERENCE 1 (bases 1 to 731)  
AUTHORS Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
TITLE A complementary view of the C.elegans genome  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
1. 731  
/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone\_lib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"

BASE COUNT 236 a 128 c 152 g 214 t 1 others  
ORIGIN

Query Match 48.4%; Score 707; DB 10; Length 731;  
Best Local Similarity 98.9%; Pred. No. 1.7e-120;  
Matches 722; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

Qy 722 aagccagcttgaaagcaattctcaagattgtaccatatcaacatctctatttcacc 781  
|||||  
Db 731 AAGGCCAGCTTTGAAAGCAATTTCTCAAGATTGTACCATATCAACATCTCTATTTCACC 672  
|||||  
Qy 782 gcaatgtctcaatgctccgtttctctatcgactggtcgagtcagttcaatgggtatcaca 841  
|||||  
Db 671 GCAATGCTTCCAATGCTCCGTTTCTCATCGAGCTGTGTCAGTCAATGGGTATTTCAAA 612  
|||||  
Qy 842 gagaatcaatggagtcacagctctatcaagaatgcatctctgggagcaagcaacaatt 901  
|||||  
Db 611 GAGAATCAATGGAGTCAAGGTCTATCAAGAAATGCTATTCTGGGCAAGCAACAATT 552  
|||||  
Qy 902 gttgacattggctgggtattctctatcaattgtcttattaccaacatggccacttcgg 961  
|||||  
Db 551 GTTGGACATTTGGCTTGGGTATCTATCAATTGTCTTATTACCAACATGGCCACTTCGG 492  
|||||  
Qy 962 gttgcttatttcattatctcaaaatgggagggagcccttttgattgctcaagttagtcaact 1021  
|||||  
Db 491 GTTGTCTATTTCATTTATTCAAATGGAGGAGGCGCTTTTGTATGCTCAGCTAGTCACT 432  
|||||  
Qy 1022 ttcaaccaataactctgttgaatgattccagccaatctcaaatctcaattttaaacaaactcgcc 1081  
|||||  
Db 431 TTCACCAATAACTCTGTGTGAATGATTCAGCCAAATTCGAAATTTTAAACAACATTCGCC 372  
|||||  
Qy 1082 gctcttcaaatatttgaccacacgaacatgactccatctccattcatgattggcttgg 1141  
|||||  
Db 371 GCTCTTCAAAATTTTGACCACACGCAACATGACTCCATCTCCATTCATTGATGGCTTGG 312  
|||||  
Qy 1142 ggtgagcaattatcatagatcgagcaccactgtgtcccaacaatgccacgttgcaatctg 1201  
|||||  
Db 311 GGTGACTCAATATTCAGATCGAGCACCACCTTGTTCGCCAACAATGCCACGTTGCAATCTG 252  
|||||  
Qy 1202 aatgctgctgaaatatgtgaagaatggtgcaagaagaataatcttcccttaacctgtc 1261  
|||||  
Db 251 AATGCTTCGATGAATATGTGAAAGAATGGTGCAAAAGAGAATAATCTTCTTACCTTCGTC 192  
|||||

Qy 1262 gatgactactttgacggatgatcaatgaatttgacaacaaatgaaataatggctgagcac 1321  
|||||  
Db 191 GATGACTACTTTGACGGATATGCAATGAATTTGCAACAATTTGAAATAATGCTGAGCAC 132  
|||||  
Qy 1322 attcaagctaaagctgcctaaacaatctcgtggtgttcaaaaaagttttttcttg--tttttt 1379  
|||||  
Db 131 ATTCAAGCTTAAAGCTGCCCTAAACAATCTGGGTGTTTCAAAAAGTTTTTTCTGTTTTTT 72  
|||||  
Qy 1380 aaatttaattctttgaaaattattgttttccgtcattctcctcattcccttctctggt 1439  
|||||  
Db 71 AAATTTAATCTTTGAATAATATTGTTTTCGCTCATCTTCCATTCCTCCATTCCTTCGAGT 12  
|||||  
Qy 1440 agaaaataaaa 1449  
|||||  
Db 11 AGAAAAATAA 2  
|||||

RESULT 13  
AU210015/c 720 bp mRNA linear EST 17-JUL-2001  
LOCUS AU210015  
DEFINITION AU210015 unpublished oligo-capped cDNA library, stage L1  
Caenorhabditis elegans cDNA clone yk751h06 3', mRNA sequence.  
ACCESSION AU210015  
VERSION AU210015.1 GI:14846014  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans.  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
REFERENCE 1 (bases 1 to 720)  
AUTHORS Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
TITLE A complementary view of the C.elegans genome  
JOURNAL Unpublished (2001)  
COMMENT Contact: Yuji Kohara  
Genome Biology Lab.  
National Institute of Genetics  
Yata 1111, Mishima, Shizuoka 411, Japan  
Tel: 81-559-81-6854  
Fax: 81-559-81-6855  
Email: ykohara@lab.nig.ac.jp.  
Location/Qualifiers  
1. 720  
/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone\_lib="yk751h06"  
/clone\_lib="unpublished oligo-capped cDNA library, stage  
L1"  
/sex="Hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"

BASE COUNT 234 a 127 c 152 g 207 t  
ORIGIN

Query Match 48.2%; Score 704.4; DB 9; Length 720;  
Best Local Similarity 99.6%; Pred. No. 5.1e-120;  
Matches 717; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 713 tgaagataaaggcagctttgaaaaagcaattctcaagattgtaccatatcaacatctc 772  
|||||  
Db 720 TGCAAGTATAAGCCAGCTTTGAAAAAGCAATTTCTCAAGATTGTACCATATCAACATCTC 661  
|||||  
Qy 773 tatttcacgcgaatgcttccaatgctccgtttctcattgagctggtcagtcagttcaatgg 832  
|||||  
Db 660 TATTTCACCGCAATGCTTCCAATGCTCGGTTTCTCATGAGCTGGTCAAGTCAATGAGT 601  
|||||  
Qy 833 gtattcaagaagaatcaaatgagtagaaggtctatcaagaagaatgcatctctgggagcaa 892  
|||||  
Db 600 GTATTTCAAGAGAATCAAAATGGAGTACAGGTCTATCAAGAAATGCATTTCTGGGAGCAA 541  
|||||  
Qy 893 gaacaattgttggacattgggctgggtattctctatcaattgttcttattaccacaatgg 952  
|||||

Db 540 GCAACAATTTGGACATTTGGGCTTGGGTATCTATCAATTTGTTCTTATTACCAACATGG 481  
Qy 953 ccactcgggtgcttattcattattccaaatggagagagcccttttgatgctcac 1012  
Db 480 CCACATTCGGGTGCTTTATTTCAATTTTCACAAATGGGAGGAGGCGCTTTGATTGCTCAC 421  
Qy 1013 gtagcactttcaaccataaactcgttgataagatgccagcgaatctcgaattttaaac 1072  
Db 420 GTAGTCACCTTCACACCAATACTGTTGATTAAGTATCCAGCCAAATCTCGAATTTTAAC 361  
Qy 1073 aactcgcgcgtctcctcaaatTTTgaccacacgcgaacatgactcctcattcattgat 1132  
Db 360 AACTTCGCGGCTCTCAAAATTTGACCACACGCCAACATGACTCCTATCTCCATTCATTGAT 301  
Qy 1133 tggcttgggtgactcaaatattcagatcgacgaccactgtgtcccaacaatgccacgt 1192  
Db 300 TGGCTTTGGGCTGACTCAATATCAGATCGAGCACCACCTTTGTTCCCAACAAATGCCAGT 241  
Qy 1193 tgcaatcgaatgcttgcgtgaaataatgtgaaagaatgtgcaagagaataatcttct 1252  
Db 240 TGCAATCTGAATGCTTGCAATCAATATGTGAAGAATGGTGCAAGAGAAATATCTTCT 181  
Qy 1253 tacctcgtcgtactcattgacggatatacgaatgaatttgcaacaattgaaataatg 1312  
Db 180 TACCTCGCGATGACTACTTTGACGGATATGCAATGAATTTGCAACAAATTTGAAATAATG 121  
Qy 1313 gctgagcacattcaagctaaagctgcctcaacaatcgtggtgttcacaaagtcttctt 1372  
Db 120 GCTGAGCACATTCAGCTAAAGCTGCCTTAAACAATCTGGGTGTTCAAAAAGTTTTTCTT 61  
Qy 1373 g--tttttaatttaattcttggaaattattgttttccgtcattcttccctccatccc 1430  
Db 60 GTTTTTTTTAAATTAATTTCTTTGAAATTAATTTGTTTTCGCTGATCTTCTCCTCCATTCCC 1

RESULT 14  
BJ151779/c

LOCUS  
DEFINITION BJ151779 unpublished oligo-capped cDNA library, C. elegans L1 stage  
Caenorhabditis elegans cDNA clone yk1299e08 3', mRNA sequence.

ACCESSION BJ151779  
VERSION BJ151779.1 GI:18319764  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans.  
ORGANISM Caenorhabditis elegans.

REFERENCE  
AUTHORS Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.

TITLE A complementary view of the C.elegans genome  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
1. .721  
/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1299e08"  
/clone\_lib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"

FEATURES  
source

1. .721  
/organism="Caenorhabditis elegans"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1299e08"  
/clone\_lib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"

BASE COUNT 236 a 128 c 149 g 207 t 1 others  
ORIGIN

Query Match 48.1% Score 702.8; DB 10; Length 721;  
Best Local Similarity 99.3%; Pred. No. 1e-119;  
Matches 716; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

Qy 710 ttgtcgaagtataaggcagcgtttgaaaaaqaattctcaagattgtaccatatcaacat 769  
Db 721 TTGTGCAAGTATAAGGCCAGCTTTGAAAAAGCAATTTCTCAAGATTTGTACCATATCAACAT 662  
Qy 770 cctatttcaaccgcaatgcttccaaatgctcogtttctcactggaactggtcagtcagtcaaa 829  
Db 661 CTCTATTTCACCGCAATGCTTCCAAATGCTCGGTTTCTCATGACTGCTGCTAGTTCAAA 602  
Qy 830 tgggtattcgaagagaatcaaatgagtcacaggtctatcaagaataatgcattctgggag 889  
Db 601 TGGGTATTCAAGAGATCAAAATGGAGTACAGGTCTATCAAGAAATATGCATTCTGGGAG 542  
Qy 890 caagcaacaattgttgacattgggcttgggtattctatcattcaattgttcttattaccaca 949  
Db 541 CAAGCAACAATTTGTGACATTTGGCTTGGGTATTTCTATCAATTTGTTCTTATTACCAACA 482  
Qy 950 tggccactcgggtgcttatttccattatttccaaatgggagagccttttgcattgct 1009  
Db 481 TGGCCACTTCGGGTGCTTATTTCAATTTTCAAAAATGGGAGGAGGCGCTTTGATTGCT 422  
Qy 1010 cactgtagcactttcaaccataaactcgttgataagtatccagcgaatctcgaatttta 1069  
Db 421 CACGTAGTCACCTTTCAACCATACTCTGTGATTAAGTATCCAGCCAAATTTCTCGAATTTTA 362  
Qy 1070 acaacttcgcgcgtcttccaaattttgaccacacgcgaacatgactcctcattccatt 1129  
Db 361 AACAACTTCGCGCTCTTCAAAATTTTGACCACACGCAACATGACTCCATCTCCATTTCATT 302  
Qy 1130 gattggcttgggtggaactcaattatcagatcgagcaccactgttcccaacaatgcc 1189  
Db 301 GATTGGCTTTGGGTGGAGCTCAATTTATCAGATCGAGCACCACCTTTGTTCCCAACAAATGCCA 242  
Qy 1190 cgttgcaatcgaatgcttgcgtgaaataatgtgaaagaatggtgcaagagaataatctt 1249  
Db 241 CGTTGCAATCTGAATGCTTGGCATGAATATGTGAAGAAATGTCGCAAGAGAAATATCTT 182  
Qy 1250 ccttacctcgtcgtactcattgacggatatacgaatgaatttgcaacaatttgaaaaat 1309  
Db 181 CTTTACCTCGTCGATGACTACTTTGACGGATATGCAATTTGTAACAATTTGAAAAAT 122  
Qy 1310 atgctgaagcacattcaagctaaagctgcctcaacaatcgtggtgttcacaaagtctt 1369  
Db 121 ATGGCTGAGCACATTCAGCTAAAGCTGCCTTAAACAATCTGGGTGTTCAAAAAGTTTTTT 62  
Qy 1370 cttg--tttttaatttaattcttggaaattattgttttccgtcattcttccctccatt 1427  
Db 61 CTTGTTTTTTTAAATTAATTTCTTTGAAATTAATTTGTTTTCGCTGATCTTCTCNCNCATT 2

RESULT 15  
BJ152662/c

LOCUS  
DEFINITION BJ152662 unpublished oligo-capped cDNA library, C. elegans L1 stage  
Caenorhabditis elegans cDNA clone yk1310b09 3', mRNA sequence.

ACCESSION BJ152662  
VERSION BJ152662.1 GI:18320647  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans.  
ORGANISM Caenorhabditis elegans.

REFERENCE  
AUTHORS Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
; Rhabditidae; Peloderinae; Caenorhabditis.  
; Rhabditidae; Chromadorea; Rhabditida; Rhabditoidea  
Eukaryota; Metazoa; Nematoda;

TITLE and Sugano, S.  
JOURNAL A complementary view of the C. elegans genome  
COMMENT Unpublished (2002)  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES Location/Qualifiers  
source 1..725  
    /organism="Caenorhabditis elegans"  
    /strain="N2"  
    /db\_xref="taxon:6239"  
    /clone="yk1310b09"  
    /clone\_lib="unpublished oligo-capped cDNA library, C.  
        elegans L1 stage"  
    /sex="hermaphrodite"  
    /tissue\_type="whole animal"  
    /dev\_stage="L1"  
BASE COUNT 235 a 131 c 149 g 210 t  
ORIGIN

Query Match 48.0%; Score 701.8; DB 10; Length 725;  
Best Local Similarity 98.9%; Pred. No. 1.5e-119;  
Matches 717; Conservative 0; Mismatches 7; Indels 1; Gaps 1;  
  
Qy 698 attccaggagatttgcaagataaagccagcttgtaaaaaagaacattctcaagattgta 757  
Db 725 ATTCCAGGAGATTGTGCAAGTATAAAGCCAGCTTTGAAAAAGCAATTCCTCAAGATTGTA 666  
  
Qy 758 ccatacaacatctctattccacgcgaatgcttcaatgctccgtttctcatgactggt 817  
Db 665 CCATATCAACATCTATTTACCGCAATGCTTCANATGCCGTTTCTCATGGACTGGT 606  
  
Qy 818 cagtcagttcaatgggtattcaagagagaatcaaatggagtagcaaggtctatcaagaata 877  
Db 605 CAGTCAGTTCAATGGGTATTCAAGAGAAATCAAAATGGAGTACAAGGCTCTATCAAGAAAT 546  
  
Qy 878 gcattctgggaggaagaacaattgttgacattggcgttggttattctatcaattgttc 937  
Db 545 GCATTCTGGGAGCAAGCAACAATTGTTGGACATTGGGCTGGGTATTCTATCAATTGTTTC 486  
  
Qy 938 ttattaccaacatggccacttcggttgcttatttcatttattcaacaatgggaggagc 997  
Db 485 TTATTACCAACATGCCACTTCGGGTGCTTATTTCATTATTTCACAAATGGGAGGAGC 426  
  
Qy 998 cttttgattgctacgttagtcaacttccaccataactcctgttgataagtatccagccaat 1057  
Db 425 CTTTTGATTGCTCAGTAGTACATTCAACCATTAACCTCTGTTGATAAGTATCCAGCCAAAT 366  
  
Qy 1058 tctcgaattttaaacaacttcgcgcgtcttcaaaatttgaccacacgcaacatgactcca 1117  
Db 365 TCTCGAATTTTAAACAACATTCGCCGCTCTTCAAAATTTTGACACACGCAACATGACTCCA 306  
  
Qy 1118 tctccattcattggtggttggtgactcaattatcatcagatcgacacacacttgctc 1177  
Db 305 TCTCATTCATGATGGCTTGGGGTGGACTCAATTAATACATCGAGCACCATCTGTTC 246  
  
Qy 1178 ccaacaatgccacgttgcaatctgaatcgttcggtgaaatattgaaagaatggtgcgaaa 1237  
Db 245 CCAACAATGCCACGTGGCAATCTGAATCTTGATGAATATGTGAAGAAATGGTGCAAA 186  
  
Qy 1238 gagaataatcttccttaacctcgtcgatgactactttgacggatatgcaatgaattgcaa 1297  
Db 185 GAGAATAATCTTCCTTACCTCGTGCATGACTACTTTGACGGATATGCAATGAATTTGCAA 126  
  
Qy 1298 caattgaaaaatatgctgagcacattcaagctaaagctgctaacaacatctgggtgttc 1357  
Db 125 CAATTGAAAAATATGGCTGTGTCGACACATTCAAGCTAAAGTCGCTAAACAACTCTGGGTGTTTC 66

Qy 1358 aaaaagtttttctcttg-ttttttaaaatttaattctttgaaaattattgttttccgctcatt 1416  
|||||  
Db 65 AAAAAAGTTTTTCTTGTTTTTTTTTAAATTAAATTCCTGAAATATTGTTTTCCGCTCATT 6  
  
Qy 1417 cttcc 1421  
|||||  
Db 5 CTTC 1

Search completed: August 7, 2002, 02:08:49  
Job time: 2895 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 16:21:46 ; Search time 31.23 Seconds  
(without alignments)  
2453.947 Million cell updates/sec

Title: US-09-555-093-2  
Perfect score: 2410  
Sequence: 1 MYVDKNASGRMKVDGKWL.....YAMNLOOLKNMAEHIOAKAA 443

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_19:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2410	100.0	443	061388	O61388 caenorhabdi
2	2407	99.9	443	023221	Q23221 caenorhabdi
3	1092.5	45.3	447	09XTB7	Q9XTB7 caenorhabdi
4	600.5	24.9	419	10 Q9SWQ9	Q9swq9 euglena gra
5	592	24.6	520	10 Q9L2M9	Q9lem9 ceratodon p
6	590.5	24.5	525	10 Q9ZNM2	Q9znw2 physcomitre
7	577	23.9	459	10 Q944W4	Q944w4 pythium irr
8	569.5	23.6	467	3 Q96VC3	Q96vc3 mucor circi
9	559	23.2	483	10 Q9LENO	Q9len0 ceratodon p
10	557	23.1	457	3 Q9UVV3	Q9uvv3 mortierella
11	556	23.1	457	3 Q9UVV3	Q9uvv3 mortierella
12	549	22.8	457	3 Q9HEV4	Q9hev4 mortierella
13	546	22.7	457	3 Q9HEV1	Q9hev1 mortierella
14	490.5	20.4	357	3 Q9HDF4	Q9hdf4 mortierella
15	463.5	19.2	449	10 Q82348	Q82348 arabidopsis
16	460.5	19.1	458	10 Q43469	Q43469 helianthus

17	457.5	19.0	448	10	O04353	O04353 borago offi
18	454.5	18.9	448	10	Q9SAU5	Q9sau5 borago offi
19	452	18.8	446	10	Q9ZTV9	Q9ztv9 ricinus com
20	448.5	18.6	469	10	Q9ZT08	Q9zt08 triticum ae
21	443.5	18.4	449	10	Q9ZRP8	Q9zrp8 brassica na
22	433	18.0	449	10	Q9ZRP7	Q9zrp7 arabidopsis
23	428.5	17.8	446	10	Q9FR82	Q9fr82 borago offi
24	426	17.7	523	3	Q9HDC8	Q9hdg8 mucor rouxi
25	415.5	17.2	454	13	Q98SW7	Q98sw7 oncorhynch
26	411	17.1	444	4	Q9NYX1	Q9nyx1 homo sapien
27	410	17.0	449	11	Q9JUE7	Q9jje7 mus musculu
28	408	16.9	444	4	Q96I39	Q96i39 homo sapien
29	406	16.8	444	4	O60427	O60427 homo sapien
30	406	16.8	444	4	Q96T10	Q96t10 homo sapien
31	404.5	16.8	444	13	Q9DEX7	Q9dex7 brachydanio
32	403.5	16.7	454	13	Q90ZE8	Q90ze8 oncorhynch
33	401	16.6	444	4	Q9NRP8	Q9nrp8 homo sapien
34	393.5	16.3	444	4	Q9S864	Q9s864 homo sapien
35	392	16.3	444	4	Q96SV3	Q96sv3 homo sapien
36	389.5	16.2	444	11	Q920R3	Q920r3 rattus norv
37	389.5	16.2	444	11	Q92122	Q92122 rattus norv
38	389.5	16.2	444	11	Q920R9	Q920r9 mus musculu
39	389	16.1	447	11	Q9EPV4	Q9epv4 rattus norv
40	382	15.9	447	11	Q920L1	Q920l1 mus musculu
41	380	15.8	445	4	Q9Y500	Q9y500 homo sapien
42	377.5	15.7	444	13	Q9DEX6	Q9dex6 cyprinus ca
43	356.5	14.8	445	13	Q90ZE9	Q90ze9 oreochromis
44	350.5	14.5	422	4	Q9H3G3	Q9h3g3 homo sapien
45	345	14.3	350	2	Q9F2M3	Q9f2m3 streptomyce

ALIGNMENTS

RESULT 1

O61388 PRELIMINARY: PRT: 443 AA.  
AC O61388;  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE DELTA6-FATTY-ACID-DESATURASE.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_taxid=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97268723; PubMed=9108131;  
RA Savanova O., Smith M.A., Lapinskas P., Stobart A.K., Dobson G.,  
Christie W.W., Shewry P.R., Napier J.A.;  
"Expression of a borage desaturase cDNA containing an N-terminal  
cytochrome b5 domain results in the accumulation of high levels of  
delta6-desaturated fatty acids in transgenic tobacco.";  
Proc. Natl. Acad. Sci. U.S.A. 94:4211-4216(1997).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98149727; PubMed=9480865;  
RA Napier J.A., Hey S.J., Lacey D.J., Shewry P.R.;  
"Identification of a caenorhabditis elegans Delta6-fatty-acid-  
desaturase by heterologous expression in saccharomyces cerevisiae.";  
Biochem. J. 330:611-614(1998).  
DR EMBL: AF031477; AACI5586.1;  
DR InterPro: IPR001199; Cyt\_B5.  
DR IncePro: IPK000566; Lipocin\_cytFABP.  
DR PROSITE: PS50255; CYTOCHROME\_B5\_2; 1.  
DR PROSITE: PS00213; LIPOCALIN; UNKNOWN\_1.  
SQ SEQUENCE 443 AA; 51740 MW; 9513CA7C5A7E9A06 CRC64;

April 97

Query Match 100.0%; Score 2410; DB 5; Length 443;  
Best Local Similarity 100.0%; Pred. No. 1.9e-194;  
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MVDKNASGLRMKVDGKWLILSEELVKKHPGGAVIEQYRNSDATHIFHAFHEGSSQAYKQ 60
Db 1 MVDKNASGLRMKVDGKWLILSEELVKKHPGGAVIEQYRNSDATHIFHAFHEGSSQAYKQ 60
Qy 61 LDLLKKHGEHDEFLEKQLEKRLDKVDINVSAYDVSVAQEKMKWSEFEKLRQKLDHDDGLMK 120
Db 61 LDLLKKHGEHDEFLEKQLEKRLDKVDINVSAYDVSVAQEKMKWSEFEKLRQKLDHDDGLMK 120
Qy 121 ANETYFLFKALSTLSIMAFAYLOYLWYITSACLLALAWOQFGWLTHFCHQOQPTKNRP 180
Db 121 ANETYFLFKALSTLSIMAFAYLOYLWYITSACLLALAWOQFGWLTHFCHQOQPTKNRP 180
Qy 181 LNDTISLFFGNFLOGFSRDWKKHNTHTHAATNVIDHDGDDIDLAPLFAFIPGDLCKYKAS 240
Db 181 LNDTISLFFGNFLOGFSRDWKKHNTHTHAATNVIDHDGDDIDLAPLFAFIPGDLCKYKAS 240
Qy 241 FEKAILKIVPYQHLYFTAMLPMLRFSWTGQSVQWVFKEQNMKEYKYORNAFWEQATIVGH 300
Db 241 FEKAILKIVPYQHLYFTAMLPMLRFSWTGQSVQWVFKEQNMKEYKYORNAFWEQATIVGH 300
Qy 301 WAWFYQLFLLPTWPLRVAYFIISOMGGGLLIAHVTFNHNNSVDKYKEMCKENNLPLYLVDY 420
Db 301 WAWFYQLFLLPTWPLRVAYFIISOMGGGLLIAHVTFNHNNSVDKYKEMCKENNLPLYLVDY 420
Qy 361 ILTTRNMTSPFIDWLGGLNYQIEHHLFPTMPRCNLNACVYKWKEMCKENNLPLYLVDY 420
Db 361 ILTTRNMTSPFIDWLGGLNYQIEHHLFPTMPRCNLNACVYKWKEMCKENNLPLYLVDY 420
Qy 421 FDGYAMNLQQLKNAEHIQAKAA 443
Db 421 FDGYAMNLQQLKNAEHIQAKAA 443
```

RESULT 2

```
Q23221 ID Q23221 PRELIMINARY; PRT; 443 AA.
-AC Q23221;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE W08D2.4 PROTEIN.
GN W08D2.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Swinburne J., Ainscough R.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: Z70271; CAA94233.2; -.
DR InterPro: IPR001199; CytL5.
DR InterPro: IPR000566; Lipocin_cytL5.
DR PROSITE: PS0255; CYTOCHROME_B5_2; 1.
DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_1.
SQ SEQUENCE 443 AA; 51772 MW; 9513D611ECB99A06 CRC64;
```

Query Match 99.98; Score 2407; DB 5; Length 443;  
Best Local Similarity 99.88; Pred. No. 3.4e-194;  
Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MVDKNASGLRMKVDGKWLILSEELVKKHPGGAVIEQYRNSDATHIFHAFHEGSSQAYKQ 60
Db 1 MVDKNASGLRMKVDGKWLILSEELVKKHPGGAVIEQYRNSDATHIFHAFHEGSSQAYKQ 60
```

```
Qy 61 LDLLKKHGEHDEFLEKQLEKRLDKVDINVSAYDVSVAQEKMKWSEFEKLRQKLDHDDGLMK 120
Db 61 LDLLKKHGEHDEFLEKQLEKRLDKVDINVSAYDVSVAQEKMKWSEFEKLRQKLDHDDGLMK 120
Qy 121 ANETYFLFKALSTLSIMAFAYLOYLWYITSACLLALAWOQFGWLTHFCHQOQPTKNRP 180
Db 121 ANETYFLFKALSTLSIMAFAYLOYLWYITSACLLALAWOQFGWLTHFCHQOQPTKNRP 180
Qy 181 LNDTISLFFGNFLOGFSRDWKKHNTHTHAATNVIDHDGDDIDLAPLFAFIPGDLCKYKAS 240
Db 181 LNDTISLFFGNFLOGFSRDWKKHNTHTHAATNVIDHDGDDIDLAPLFAFIPGDLCKYKAS 240
Qy 241 FEKAILKIVPYQHLYFTAMLPMLRFSWTGQSVQWVFKEQNMKEYKYORNAFWEQATIVGH 300
Db 241 FEKAILKIVPYQHLYFTAMLPMLRFSWTGQSVQWVFKEQNMKEYKYORNAFWEQATIVGH 300
Qy 301 WAWFYQLFLLPTWPLRVAYFIISOMGGGLLIAHVTFNHNNSVDKYKEMCKENNLPLYLVDY 420
Db 301 WAWFYQLFLLPTWPLRVAYFIISOMGGGLLIAHVTFNHNNSVDKYKEMCKENNLPLYLVDY 420
Qy 361 ILTTRNMTSPFIDWLGGLNYQIEHHLFPTMPRCNLNACVYKWKEMCKENNLPLYLVDY 420
Db 361 ILTTRNMTSPFIDWLGGLNYQIEHHLFPTMPRCNLNACVYKWKEMCKENNLPLYLVDY 420
Qy 421 FDGYAMNLQQLKNAEHIQAKAA 443
Db 421 FDGYAMNLQQLKNAEHIQAKAA 443
```

RESULT 3

```
Q9XTB7 ID Q9XTB7 PRELIMINARY; PRT; 447 AA.
AC Q9XTB7;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE DELTA5-FATTY ACID DESATURASE.
GN FAT-4 OR DES-5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99117288; PubMed=9917342;
RA Watts J.L., Browse J.;
RT "Isolation and characterisation of a delta5-fatty acid desaturase from
RT Caenorhabditis elegans."
RL Arch. Biochem. Biophys. 362:175-182(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99059458; PubMed=9845325;
RA Michaelson L.V., Napier J.A., Lewis M., Griffiths G., Lazarus C.M.,
RA Stobart A.K.;
RT "Functional identification of a fatty acid delta5 desaturase gene from
RT Caenorhabditis elegans."
RL FEBS Lett. 439:215-218(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Swinburne J.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: AF114440; AAD13294.1; -.
DR EMBL: AF078796; AAC95143.1; -.
DR EMBL: Z81122; CAB61031.1; -.
SQ SEQUENCE 447 AA; 52348 MW; D7E09CE0E2975015 CRC64;
```



```
Db 206 KTLINVSIVATSI AIIISLYKSYRA-VLLSASLMGLFIQQCGWLSHDFLHHQVFEETR WIND 264
Qy 184 TISLFFNGFLOGGSRWKKDKHNTHTHAATNVIDH-----DGDIDLAPLEAFIPGDLCKYK 238
Db 265 VGVVGNVVLGFSVSNWKKYKHNJLHHAAPNECDQKYPIDEDIDTLPILIAWSKDLLATVE 324
Qy 239 ASFEKAILKIVPYQHLVFTAMLPMLRFSWTSQSVQWVFK-ENQMEYKVVYORNAFWEQATI 297
Db 325 S---KTLMLRVLYQHLFELVLLTPARASWLFWSAAFTLRPELTIGEXLLER-----GTM 375
Qy 298 VGHWA-----VFYQLFLLPTWPLRVAYFIISOMGGGLLIAHVVTNNNSVDKYPANRIL 353
Db 376 ALHYIWFNSVAFY---LLPGWK-PVWVWVSELMSGFLGTVFVLSHNGMEVYNTS---- 427
Qy 354 NFAALQILTRNMTSPFIDWGLGNYQIEHHLFPTMPRCNLNACVKYKWKCKENNL 413
Db 428 KDFYNAQIASTRDITKAGVNDWFTGGNROIEHHLFPTMPRNLNKLISPHVETLCKRAHGL 487
Qy 414 PYLVDDYFDGYAMNLOOLKNMAE---HIQAKAA 443
Db 488 VYEDVSMASGYRVKLTLOKOVADAASHQOLAAS 520

RESULT 6
Q92NW2 PRELIMINARY; PRT; 525 AA.
AC Q92NW2;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE DELTA6-ACYL-LIPID DESATURASE.
GN DES6.
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
OX NCBI_TaxID=3218;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HEDW. B. S. G.;
RX MEDLINE=98416756; PubMed=9744093;
RA Girke J., Schmidt H., Zaehnering U., Reski R., Heinz E.;
RT "Identification of a novel delta 6-acyl-group desaturase by targeted
RT gene disruption in Physcomitrella patens.";
RC Plant J. 15:39-48(1998).
CL -1- SIMILARITY: BELONGS TO THE CYTOCHROME B5 FAMILY.
DR EMBL; AJ222981; CAAL1033.1; -
DR EMBL; AJ222980; CAAL1032.1; -
DR HSSP; P04166; IB5M.
DR InterPro; IPR001199; Cyt_B5.
DR Pfam; PF001173; heme_1; 1.
DR PROSITE; PS50255; CYTOCHROME_B5_2; 1.
KW Heme.
SQ SEQUENCE 525 AA; 59369 MW; 530F158B0C97C83F CRC64;
```

```
Query Match 24.5%; Score 590.5; DB 10; Length 525;
Best Local Similarity 32.5%; Pred. No. 2.6e-41;
Matches 140; Conservative 76; Mismatches 160; Indels 55; Gaps 12;

Qy 19 LYLSEELVKKPGGAVIEQYRNSDATHIFHAFHEGSSQAYKQDLKLLKHGHDFLEKQL 78
Db 127 VYDVSNAFDEHPGGSVISTYFGRDGTDFVSSFHAASHTWKILQ-----DFYIGDV 175
Qy 79 EKRLDKVDINVSAYDVSAQEKKVESFEKLRQKLIHDDGLMKNATYFLFKAI STLSTMA 138
Db 176 ER-----VEPTPELKLDFREMRALFLREQLFKSKLYIVMKLLTNVAL-- 218
Qy 139 FAFYLYQLGW-----YTTSACLLALAAWQFGWLTHFEFCHQOQTKNRPNDTISLFFGNE 192
Db 219 FAASATIAICWSKTISAVLASACMAALCFQCGWLSHDFLHNVQVETRWLNEVGVVIGNA 278
Qy 193 LQGSRWKKDKHNTHTHAATNVIDH-----DGDIDLAPLEAFIPGDLCKYKASFEKAILK 247
```

```
Db 279 VLGFSTGWWKEKKNLHHAAPNECDQTYQPIDEDIDTLPILIAWSKDLIATVE---NKTFLR 335
Qy 248 IVPYQHLVFTAMLPMLRFSWTSQSVQWVFKENQMEYKVVYORNAFWEQATIVGHWA--VFY 306
Db 336 ILQYQHLFFNGLLFFARGSWLFW--WRYTSTAVLSPV---DRLLEKGTVLPHYFWFVGT 390
Qy 307 QLFLLLPTW-PLRVAYFIISOMGGGLLIAHVVTNNNSVDKYPANRILNFAALQILTR 365
Db 391 ACYLLPGWKPL--VWMAVTELMGMLLGFVLSHNGMEVYNTS---KEFVSAQIVSTR 444
Qy 366 NMTSPSPIDMLWGLNLYQIEHHLFPTMPRCNLNACVKYKWKCKENNLPLYLVDDYFDGYA 425
Db 445 DIKGINFNDWFTGGNROIEHHLFPTMPRNLNKLIAPRVEFCKKHGLVEDVSIATGTC 504
Qy 426 MNLOOLKNMAE 436
Db 505 KVLKALKEVAE 515

RESULT 7
Q944W4 PRELIMINARY; PRT; 459 AA.
AC Q944W4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE DELTA-6 FATTY ACID DESATURASE.
OS Pythium irregulare.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Pythium.
OX NCBI_TaxID=36331;
RN [1]
RP SEQUENCE FROM N.A.
RA Hong H., Datta N., Mackenzie S.L., Qiu X.;
RT "A delta-6 fatty acid desaturase from Pythium irregulare.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF419296; AALI3310.1; -
SQ SEQUENCE 459 AA; 52498 MW; 286464DA761BE055 CRC64;
```

```
Query Match 23.9%; Score 577; DB 10; Length 459;
Best Local Similarity 33.9%; Pred. No. 2.9e-40;
Matches 147; Conservative 67; Mismatches 158; Indels 62; Gaps 16;

Qy 17 KWLVLSEELVKKPGGAVIEQYRNSDATHIFHAFHEGSSQAYKQDLKLLKH--GEHDEFL 74
Db 39 KW-----DSHPGGSVMLTQAGEDATDAFAVHPSSA-----LKLEQFYVGDVDETS 85
Qy 75 EKQLEKRLDKVDINVSAYDVSAQEKKM---VESFEKLRQKLIHDDGLMKNATYFLFKAI 131
Db 86 KAEIE-----GEPASDEERARRERINEFIASRYRLRVKVMGMDLYDASALYAWKL 137
Qy 132 STLSTI---MAFAYLOYLWYITSACLLALAWQFGWLTHFEFCHQOQTKNRPNDTISL 187
Db 138 STFGIAVLSMAICFFNFNSFAMVYVAGVIMGLFYQOSWLAHDFLHNOVCENRTLGNLIGC 197
Qy 188 FFGNFLAGSRDWMKKHNTHTHAATNVIDH-----DGDIDLAPLEAFIPGDLCKYK 238
Db 198 LVGNAGWGFSGVQWKKKHNHLHHAHPNL--HSAKDEGFIGDPIIDTMTPLLAWSKE---MAR 252
Qy 239 ASFEKA--ILKIVPYQHLVFTAMLPMLRFSWTSQSVQWVFKENQMEYKVVYORNAF--WEQ 294
Db 253 KAFESAHPGFIRNOAFLYPFLLL-LARLSWAQSFFYVFTE--FSFGIFDKVDFDGEK 309
Qy 295 ATIVGHWA-----VFYQLFLLPTWPLRVAYFIISOMGGGLLIAHVVTNNNSVDKYPAN 349
Db 310 AGLVHIWQIAIPIFCNMSLFE---GVAYFLMGQSCGLLLALVFSIGHNGSVYERE 365
Qy 350 SRLNNFAALQILTRNMTSPFIDWGLGNYQIEHHLFPTMPRCNLNACVKYKWKCK 409
Db 366 TK--PDFWOLQVTTTRINRASVFMDFWTGGNLYQIDHHLFPLVPRHNLPKVNVLIKSICK 423
Qy 410 ENNLPLYLVDDYFDG 423
```

```
Db 424 EFDIPFHTGFWEQ 437
RESULT 8
Q96VC3 PRELIMINARY; PRT; 467 AA.
AC Q96VC3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DELTA-6 FATTY ACID DESATURASE.
GN D6D.
OS Mucor circinelloides.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Mucor.
OX NCBI_TaxID=36080;
RN [1]
RP SEQUENCE FROM N.A.
RA Michinaka Y., Aki T., Murashima C., Shimauchi T., Nakajima T.,
RA Kawamoto S., Shigeta S., Ono K., Suzuki O.;
RT "Cloning and characterization of a delta-6 fatty acid desaturase from
RT Mucor circinelloides.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB052086; BAB69055.1; -.
SQ SEQUENCE 467 AA; 53387 MW; 8162163D30CB5C591 CRC64;

Query Match 23.68; Score 569.5; DB 3; Length 467;
Best Local Similarity 32.78; Pred. No. 1.3e-39;
Matches 152; Conservative 69; Mismatches 191; Indels 53; Gaps 15;

Qy 2 VVDKNASGRMKVGVKWLILSEELVKKHPGG-AVIEQYRNSDATHIFHAFHEGSSQAYKQ 60
Db 25 VLDKKPEARKLIVENKVYDITDFVDHPGGERVLLTQEGDADTVHEMPPS--AY-- 80
Qy 61 LDLLKKGHEHDFLEKOLEKRL-----DKVDINVSAYDSVAQEKMKVSEFKLRQKLDH 115
Db 81 -ELLANCYVGD-----CEPKLPIDSTDKKALNSAAF-----AQE-----TRDLRDKLEK 123
Qy 116 DGLMKANETFLFKAISTLSMAFAFVLYQLGW-----YITSACLLALAWQFGWLTH 168
Db 124 QGYFDASTGYFYKYSTLLVCIVGLAI-LKAWGRESTLAVFIAASLVGLFWQCGWLIAH 182
Qy 169 EFCHQOQTKNRPLNDTISLFFGNFLOGFSRDWKKDKHNTHTHAATNVIDHDGIDILAPLF- 227
Db 183 DYAHYQVTKDPNVNLFVTFGNLVQGSLSWKKKHNTHASTNVSGEDPDIDTAPILL 242
Qy 228 -----AFIPGDLCYKASFEKAILK-IVPYQHLYFTAMPLRFSWTGQSVQVWFRENQ 280
Db 243 WDEFANVFYXGLKDNASGDFRFAEHLPTQTRYFFILGFARTSWAIQSIYSFKNET 302
Qy 281 MEYKVYORNAFWEQATVGHAWVYQILFLPTWPLRV-----AYFILLSQMGGLLIAHV 336
Db 303 LNKs--KLLSWCERFLIVHWFFTYCTI---AWISSIRNIAEFFVVSQITGYLLAIVF 357
Qy 337 TFNHSVDKYPANSRIINFAALQILTRNMTSPFIDWLMGGLNYQIEHHLFPTMPRCN 396
Db 358 AMNHGMPVYSPEANHTFEVLCITGROVNCIVFGDWMGGLNYQIEHHLFPFEMPRHH 417
Qy 397 LNACVKYKWKCKENNLPLYVDDVDFGVMNLQOLKNMAEHQAK 441
Db 418 LSKVKSMVKPIAQYNIPY-HDTTVIGTIEVLQTLDFVQKISQK 461

RESULT 9
Q9LENO PRELIMINARY; PRT; 483 AA.
AC Q9LENO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DELTA 6-FATTY ACETYLENASE.
```

```
OS Ceratodon purpureus (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Dicranidae; Dicranales; Ditrichaceae; Ceratodon.
OX NCBI_TaxID=3225;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT3; TISSUE=PROTONEMATA;
RX MEDLINE=20307617; PubMed=10848999;
RA Sperling P., Lee M., Girke T., Zaehring U., Stymne S., Heinz E.;
RT "A bifunctional delta 6-fatty acyl acetylenase/desaturase from the
RT moss Ceratodon purpureus. A new member of the cytochrome b5
RT superfamily.";
RL Eur. J. Biochem. 267:3801-3811(2000).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B5 FAMILY.
DR EMBL; AJ250734; CAB94992.1; -.
DR HSSP; P04166; 1B5M.
DR InterPro; IPR001199; Cyt_B5.
DR Pfam; PF00173; heme_1; 1.
DR PROSITE; PS50255; CYTOCHROME_B5_2; 1.
KW Heme.
SQ SEQUENCE 483 AA; 54857 MW; C451D042169AB1C2 CRC64;

Query Match 23.28; Score 559; DB 10; Length 483;
Best Local Similarity 32.68; Pred. No. 1e-38;
Matches 136; Conservative 66; Mismatches 155; Indels 60; Gaps 13;

Qy 18 WLYLSEEL-----VKKHPGGAVIEQYRNSDATHIFHAFHEGSSQAYKOLDLLKKHGEHD 71
Db 78 WMIVKEKYDISRADDHPGGTVISTYFGRDGTDFATFHPPA--AWKQL-----ND 127
Qy 72 EFL-EKQLEKRLKVDINVSAYDSVAQEKMKVSEFKLRQKLDHDDGLMKANETVFLFK- 129
Db 128 YYIGDLAREEPLD-----ELLKDYRDMRAEFVREGLFKSKAWFLLOT 170
Qy 130 ----AISTLSMAFAFYLOYLGYIT-SACLLALAWQFGWLTHFFCHQOQTKNRPLNDT 184
Db 171 LINAALFAASTATICYDKSY--WAIVLSASLGMFLVQCGWLHDLFHQQVFENNTANSF 228
Qy 185 ISLFFGNFLOGFSRDWKKDKHNTHTHAATNVIDH-----DGDIDILAPLAFIPGDLCYKA 239
Db 229 FGYLEGNCVLFGSVSWRTKKNIHTAPNECDEQYTPLEDEDITLPIANSKEILATVES 288
Qy 240 SFEKAILKIVPYQHLYFTAMPLRFSWTGQSVQVWFKENQMEYKVYORNAFWEQATIVG 299
Db 289 ---KRILRVLYQHYVMILPFLFMARYSWTFGSLTFTFNPDLSTTK-----GLIEKGTVA 340
Qy 300 HWAWVYOLF-LLPTWPLRVAYFIISOMGGILLTAHVVTFNHNSVDKYPANSRIINFAA 358
Db 341 HYAWFSWAAPHILPGVAKPLAMVATELVAGLLLGFTLSHGKVEYNES----KDFVR 396
Qy 359 LOILTRNMTSPFIDWLMGGLNYQIEHHLFPTMPRCNLACVKYKWKCKENNLPLY 415
Db 397 AOVIITRNKGNWFNDWFTGGDLTQIEHHLFPTMPRHNYPKIAPOVEALCKKHGLE 453

RESULT 10
Q9UVV3 PRELIMINARY; PRT; 457 AA.
AC Q9UVV3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DELTA-6 FATTY ACID DESATURASE.
OS Mortierella alpina.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mortierellaceae;
OC Mortierella.
OX NCBI_TaxID=64518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC32221;
RX MEDLINE=99406036; PubMed=10478922;
RA Huang Y.S., Chaudhary S., Thurmond J.M., Bobik E.G. Jr., Yuan L.,
```

RA Chan G.M., Kirchner S.J., Mukerji P., Knutzen D.S.;  
RT \*Cloning of delta12- and delta6-desaturases from Mortierella alpina  
RT and recombinant production of gamma-linolenic acid in Saccharomyces  
RT cerevisiae\*;  
RL Lipids 34:649-659(1999).  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B5 FAMILY.  
DR EMBL: AF110510; AAF08685.1; -.  
DR HSSP: P04166; 1B5M.  
DR InterPro: IPR001199; Cyt\_B5.  
DR Pfam: PF00173; heme\_1.1.  
DR PROSITE: PS00191; CYTOCHROME\_B5\_1; UNKNOWN\_1.  
DR PROSITE; PS0255; CYTOCHROME\_B5\_2; 1.  
KW Heme.  
SQ SEQUENCE 457 AA; 51837 MW; D9016986911450A CRC64;

Query Match 23.1%; Score 557; DB 3; Length 457;  
Best Local Similarity 32.1%; Pred. No. 1.4e-38;  
Matches 143; Conservative 64; Mismatches 187; Indels 52; Gaps 11;

QY 12 MKVDGKWLSEELVKKHGGAVIEQYRNSDATHIFHAFHEGSSQAYKQLDLLKKHGEHD 71  
DB 34 MIIDNK-VYDREVPDPHGGSVILTHVGKDGTDVDTFHPAAW-----77  
QY 72 EFLEKQLEKRLDKVDINVSAYDVSAQEKKVESFEKLRQLKLDHDDGLMKANETVFLFKAI 131  
DB 78 ETLANFVVGDI DESDRDIKNDDFA-AEVRKLRTLFQSL-----GYDSSKAYIAFKVS 129  
QY 132 STLSIMAFAYLYQLGW-----YITSACLLALAWQFCWLTHFCHQOQPTKNRPLNDT 184  
DB 130 FNLCINGLSTVI-VAKWGQSTLANVLSAALLGLFWQCGWLADFLHHQVFDQRFWGD 188  
QY 185 ISLFFGFLGSDRWKDKHNTTHAATNVHDHGDIDLAPL-----FAFIG-D 233  
DB 130 FNLCINGLSTVI-VAKWGQSTLANVLSAALLGLFWQCGWLADFLHHQVFDQRFWGD 188  
QY 185 ISLFFGFLGSDRWKDKHNTTHAATNVHDHGDIDLAPL-----FAFIG-D 233  
DB 189 FGAFGLGVCQGFSSWKKDKHNTTHAATNVHDHGDIDLAPL-----FAFIG-D 248  
QY 234 LCKYKASFEKAILKIVPYQHLYFTAMLPMLRFSWTGQSVQVFKENQMEYKVYQ--NAF 291  
DB 189 FGAFGLGVCQGFSSWKKDKHNTTHAATNVHDHGDIDLAPL-----FAFIG-D 248  
QY 234 LCKYKASFEKAILKIVPYQHLYFTAMLPMLRFSWTGQSVQVFKENQMEYKVYQ--NAF 291  
DB 249 LTRMWSRF-----MVLNQTFWFPILSFARLSQILFLPQGAHPSGARVPISL 302  
QY 292 WEQATIVGHAWFYQLFLPTWPLR-VAYFIISQMGGLLIARHVTFNHNSVDKYPANS 350  
DB 292 WEQATIVGHAWFYQLFLPTWPLR-VAYFIISQMGGLLIARHVTFNHNSVDKYPANS 350  
DB 303 VEQLSLAMHTWTYLATMFLFKDPVNMVYFLVSQAVCGNLLAIVFLSNHNGMPVISKEE 362  
QY 351 RIILNFAALQILTRNTPSPFFIDWLWGGLNYQIEHLLFTMPRCNLNACVKYKWKCKE 410  
DB 351 RIILNFAALQILTRNTPSPFFIDWLWGGLNYQIEHLLFTMPRCNLNACVKYKWKCKE 410  
DB 363 AVDMDFTKQIIGRDVHPGLFANWFTGGLNYQIEHLLFSPMRHNFPSKIQPAVETLCKK 422  
QY 411 NNLPLYVDYDFDGYAMNLOOLKNMAE 436  
DB 423 YNRYHTTGMIEGTAEVFSRLNEVSK 448

RESULT 11  
Q9UYV3 ID Q9UYV3 PRELIMINARY; PRT; 457 AA.  
AC Q9UYV3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE DELTA-6 FATTY ACID DESATURASE.  
OS Mortierella alpina.  
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mortierellaceae;  
OC Mortierella  
OX NCBI\_TaxID=64518;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-IS-4;  
RX MEDLINE-20035749; PubMed=10570972;  
RA Sakuradani E., Kobayashi M., Shimizu S.;  
RT "Delta 6-Fatty acid desaturase from an arachidonic acid-producing  
RT Mortierella fungus. Gene cloning and its heterologous expression in a  
RT fungus, Aspergillus.";

RL Gene 238:445-453(1999).  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B5 FAMILY.  
DR EMBL: AB020032; BAA85588.1; -.  
DR HSSP: P04166; 1B5M.  
DR InterPro: IPR001199; Cyt\_B5.  
DR Pfam: PF00173; heme\_1.1.  
DR PROSITE; PS00191; CYTOCHROME\_B5\_1; UNKNOWN\_1.  
DR PROSITE; PS0255; CYTOCHROME\_B5\_2; 1.  
KW Heme.  
SQ SEQUENCE 457 AA; 51816 MW; 5C4B3D7312439543 CRC64;

Query Match 23.1%; Score 556; DB 3; Length 457;  
Best Local Similarity 31.8%; Pred. No. 1.7e-38;  
Matches 142; Conservative 64; Mismatches 188; Indels 52; Gaps 11;

QY 12 MKVDGKWLSEELVKKHGGAVIEQYRNSDATHIFHAFHEGSSQAYKQLDLLKKHGEHD 71  
DB 34 MIIDNK-VYDREVPDPHGGSVILTHVGKDGTDVDTFHPAAW-----77  
QY 72 EFLEKQLEKRLDKVDINVSAYDVSAQEKKVESFEKLRQLKLDHDDGLMKANETVFLFKAI 131  
DB 78 ETLANFVVGDI DESDRDIKNDDFA-AEVRKLRTLFQSL-----GYDSSKAYIAFKVS 129  
QY 132 STLSIMAFAYLYQLGW-----YITSACLLALAWQFCWLTHFCHQOQPTKNRPLNDT 184  
DB 130 FNLCINGLSTVI-VAKWGQSTLANVLSAALLGLFWQCGWLADFLHHQVFDQRFWGD 188  
QY 185 ISLFFGFLGSDRWKDKHNTTHAATNVHDHGDIDLAPL-----FAFIG-D 233  
DB 189 FGAFGLGVCQGFSSWKKDKHNTTHAATNVHDHGDIDLAPL-----FAFIG-D 248  
QY 234 LCKYKASFEKAILKIVPYQHLYFTAMLPMLRFSWTGQSVQVFKENQMEYKVYQ--NAF 291  
DB 249 LTRMWSRF-----MVLNQTFWFPILSFARLSQILFLPQGAHPSGARVPISL 302  
QY 292 WEQATIVGHAWFYQLFLPTWPLR-VAYFIISQMGGLLIARHVTFNHNSVDKYPANS 350  
DB 292 WEQATIVGHAWFYQLFLPTWPLR-VAYFIISQMGGLLIARHVTFNHNSVDKYPANS 350  
DB 303 VEQLSLAMHTWTYLATMFLFKDPVNMVYFLVSQAVCGNLLAIVFLSNHNGMPVISKEE 362  
QY 351 RIILNFAALQILTRNTPSPFFIDWLWGGLNYQIEHLLFTMPRCNLNACVKYKWKCKE 410  
DB 351 RIILNFAALQILTRNTPSPFFIDWLWGGLNYQIEHLLFTMPRCNLNACVKYKWKCKE 410  
DB 363 AVDMDFTKQIIGRDVHPGLFANWFTGGLNYQIEHLLFSPMRHNFPSKIQPAVETLCKK 422  
QY 411 NNLPLYVDYDFDGYAMNLOOLKNMAE 436  
DB 423 YNRYHTTGMIEGTAEVFSRLNEVSK 448

RESULT 12  
Q9HEY4 ID Q9HEY4 PRELIMINARY; PRT; 457 AA.  
AC Q9HEY4;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE DELTA6-FATTY ACID DESATURASE.  
OS Mortierella isabellina.  
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mortierellaceae;  
OC Mortierella  
OX NCBI\_TaxID=91625;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ming-Chun L., Li L., Guo-Wu H., Li Z., Lai-Jun X.;  
RT "Cloning and sequencing analysis of delta6-fatty acid desaturase gene  
RT from Mortierella isabellina.";  
RL Junwu Xitong 0:0-0(2001).  
DR EMBL: AF306634; AAG38104.1; -.  
DR HSSP: P04166; 1B5M.  
DR InterPro: IPR001199; Cyt\_B5.  
DR PROSITE; PS00191; CYTOCHROME\_B5\_1; UNKNOWN\_1.  
DR PROSITE; PS0255; CYTOCHROME\_B5\_2; 1.  
SQ SEQUENCE 457 AA; 51772 MW; 868E7EB21172D5AF CRC64;

Query Match	22.88;	Score 549;	DB 3;	Length 457;
Best Local Similarity	31.88;	Pred. No. 6.6e-38;		
Matches 142;	Conservative	189;	Indels 52;	Gaps 11;

  

Qy	12	MKVDPKWLYLEELVKHPGGAVTEQVRNSDATHIFAHFEGSSQAYKQLDLLKKHGEHD	71
Db	34	MIIDNK-VYDVREVPDPHGGSVILTHVGKDGTDVDFTFHPEAAW-----	77
Qy	72	EFLEKQLEKRLKVDINVSAYDVSVQAQKKMVSEFKLRQLKLDHDDGLMKANETVFLFKAI	131
Db	78	ETLANFYVGDIDESDRAIKNDFA-AEVRKLTTLFQSL-----GYDSSKAYYAPKVS	129
Qy	132	STLSIMAFAYLYLGW-----YITSACLLALAWQOFGHLTHEFCHQOQTKNRPNDT	184
Db	130	FNLCIWLSTFI-VAKWGOTSTLANVLSAAILGLFVQOQCGWLADHDLHQVQDFRFGWL	188
Qy	185	ISLFFGNFLOQFSRDWKKDKHNTHAANVTDHGDIDLAPL-----FAFIPQ-D	233
Db	189	FGAFLGGVCQFSSWKKDKHNTHAAPNVGEGDDIDTHPELLTWSEHALEMFSDVPDEE	248
Qy	234	LCYKASFEKAILKIVPYQHLYFTAMLPLRFSGWTGQSVQVWFKENQMEYKVYQR--NAF	291
Db	249	LTRWMSRF-----WVLNQTWYFPIILSFARLSCLOSILLVLPNGQAHKPSGARVSISL	302
Qy	292	WEQATIVGHWAWVYQQLFTPLTWELR-VAYFIISOMGGGLLIAHVVTNNHNSVKYPANS	350
Db	303	VEQLSLAMHWTWYLATWFLFTKDPVNMVMVYFLVSQAVCGNLLAIVFLSNHNGMPVISKEE	362
Qy	351	RILNFAALQILTRNMTSPFIDWLGGLNAYQTEHLHFTPTMPCNLNACVYKVKEMCKE	410
Db	363	AVDMDFTKIITGRDVHPGLFANWTEGGLNAYQTEHLHFTPTMPCNLNACVYKVKEMCKE	422
Qy	411	NNPLVLYDDYDGYAMNQLQKNMAE	436
Db	423	YGVRYHTGTMGTEGTAEVFSRLNEYSK	448

  

RESULT 13				
Q9HEY1				
ID	Q9HEY1	PRELIMINARY;	PRT;	457 AA.
AC	Q9HEY1;			
DT	01-MAR-2001	(TrEMBLrel. 16, Created)		
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)		
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)		
DE	DELTA 6-FATTY ACID DESATURASE.			
OS	Mortierella alpina.			
OC	Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mortierellaceae;			
OC	Mortierella.			
OX	NCBI_TaxID=64518;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RP	Xing L., Li M., Liu L., Hu G., Zhang L.;			
RT	"Expression of Mortierella alpina delta 6-fatty acid desaturase gene			
RT	in Saccharomyces cerevisiae.";			
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.			
RL	EMBL; AF307940; AAG45092.1; -			
DR	HSSP; P04166; 1B5M.			
DR	InterPro; IPR001199; Cyt.B5.			
DR	PROSITE; PS00191; CYTOCHROME_B5_1; UNKNOWN_1.			
DR	PROSITE; PS50255; CYTOCHROME_B5_2; 1.			
SQ	SEQUENCE 457 AA; 51885 MW; 4B0AC0DD15D19DA1 CRC64;			

```

Query Match      22.7%; Score 546; DB 3; Length 457;
Best Local Similarity 32.4%; Pred. No. 1.2e-37;
Matches 144; Conservative 62; Mismatches 189; Indels 50; Gaps 11;

QY 12 MKVDGKWLYLEELVKRHPGGAVIEQYRNSDATIFHFHFEGSSOAYKQLDLLKKHGEHD 71
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
34 MIDNK-VYDVRVPEHPHGGVSVILTHVGKDGTDVDFTHFPEAAW----- 77
Db

```

Qy	72	EFLEKLEKRLDKVDINVSAYDVSAQEKKWVSEFKLRQKLHDDGLMKANETYFLPKAI	131
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	78	ETLANFYVGDDIDESDRAIKND DFA-AEYRKLRITLFQSL-----GYDSSKAYIAFKVS	129
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Qy	132	STLSIMAFPA-FYLOVLGWYIT-----SACLLALAWAQRFGLTWFHCFHQOQPKNRPLNDTI	185
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	130	FNLCIWGLSTFIVAKRGQTSTLANELSAALLGLFWQQRGLAHDLFHQVFDQRFWGDLF	189
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Qy	186	SLFNGFNLOGFSRDWKKDKHNTTHAAATVVDHDDGIDLIAPL-----FATIPG-DL	234
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	190	GAFLGGDCQGFSSSSWKKDKHNTTHAAAPNVHGEDPDIDTFLPLLTWSEHALEMFSDVDEEL	249
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Qy	235	CKYKASEEKAILKIVPYOHLVFTAMPLMLRFSWTCOSVQWVFKENOMEYKVVQR--NAFW	292
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	250	TRMWSRF-----MVLNGTWTFYFPLSFARUSCLQSLFLVLPNCQAHKPSGARVPISLV	303
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Qy	293	EQATIVGHAWVFYQLFLLTPWPLR-VAYFIISQMGGLLIAHVYTFNHNNSVDKYPANSR	351
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	304	EQLSLAHMTWYLATMELFIKDPVNMVYFVLVSQAQVGNLLAIVFSLNHNMGMPVISKEEA	363
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Qy	352	ILNFPAAALQIITTRNMTSPSIDMLWGLNGLQVIEHHLEFPTMPCRNLCACVYVKEWCKEN	411
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	364	VDMOFETKQIITGRDVBHPLGFANFTGLNFTQIEHHLEFSPMRHNFSKIQPAVETLCKKY	423
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Qy	412	NLPYLVDVDFDGYAMNLQOLKNMAE	436
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	424	GVRHTTGMIEGTAEVFSRLNEVSK	448
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
RESULT	14		
Q9HDF4			
ID	Q9HDF4	PRELIMINARY;	357 AA.
AC	Q9HDF4		
DT	01-MAR-2001	(TrEMBLrel. 16, Created)	
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)	
DE	DELTA 6-FATTY ACID DESATURASE (FRAGMENT).		
OS	Mortierella alpina.		
OC	Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mortierellaceae;		
OC	Mortierella.		
OX	NCBI_Taxid=64518;		
{1}			
SEQUENCE FROM N.A.			
RP	SEQUENCE FROM N.A.		
RP	Xing L., Li M., Liu L., Hu G., Zhang L.;		
RT	"Cloning and sequence analysis of the conserved region of delta 6-		
RT	fatty acid desaturase gene from Mortierella alpina.;"		
RT	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AF307942; AAG45094.1; -		
DR	EMBL; AF307941; AAG45093.1; -		
DR	HSSP; P04166; 1B5M.		
DR	InterPro; IPR001199; Cyt_B5.		
DR	PROSITE; PS00191; CYTOCHROME_B5.1; UNKNOWN.1.		
DR	PROSITE; PS50255; CYTOCHROME_B5_2; 1.		
FT	NON_TER	1	
FT	NON_TER	357	357
SQ	SEQUENCE	357 AA; 40796 MW; C8D9CE1283BB16B8	CRC64;

	Query Match	20.4%	Score 490.5;	DB 3;	Length 357;
	Best Local Similarity	33.2%;	Pred. No. 3.9e-33;		
	Matches 128;	Conservative	50;	Mismatches 158;	Indels 49; Gaps 10;
Qy	26 VKKHGGGAVIQYRNSDATHIFAFHGSSQAQYQLDLKKHGHDEFLKLEKRLLDKV	85	:   :		
Dd	2 VPDHPGGSVILTHGVKDCTVDFTFHPEAAW-----ETLANFYVGDIDES	46	:   :		
Oy	86 DINYSAYDVSAQEKKMVESEKIRLOKLHDDGLMKANETYELFKRAISTLSINATA-FYLQ	144			





---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 16:18:46 ; Search time 20.87 Seconds  
(without alignments)  
2039.654 Million cell updates/sec

Title: US-09-555-093-2  
Perfect score: 2410  
Sequence: 1 MYVDKNASGLRMKVDGKNLY.....YAMNLOQLKNMAEHIOAKAA 443

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71.\*

- 1: pir1.\*
- 2: pir2.\*
- 3: pir3.\*
- 4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2382	98.8	473	T26280	linoleoyl-CoA desaturase
2	1092.5	45.3	447	T43319	Delta6 fatty acid
3	1079	44.8	454	H88791	protein Tl3f2.1 [i
4	463.5	19.2	449	T26280	hypothetical prote
5	460.5	19.1	458	S68358	Delta8 sphingolipi
6	443.5	18.4	449	T50555	delta-8 sphingolip
7	433	18.0	449	T47950	delta-8 sphingolip
8	421	17.5	523	JC7556	linoleoyl-CoA desaturase
9	393.5	16.3	444	Tl3155	linoleoyl-CoA desaturase
10	389.5	16.2	444	JG0180	Delta6 fatty acid
11	296.5	12.3	345	T36617	probable Delta6 fa
12	225	9.3	368	S54809	linoleoyl-CoA desaturase
13	215.5	8.9	359	S35157	Delta6 fatty acid
14	174.5	7.2	443	T08136	probable omega-6 d
15	165.5	6.9	448	D85362	hypothetical prote
16	158.5	6.6	447	S53309	n-6 fatty acid desaturase
17	157.5	6.5	350	S43772	phosphatidylcholin
18	157.5	6.5	350	AH2005	phosphatidylcholin
19	148	6.1	424	T07742	omega-6 desaturase
20	147.5	6.1	424	JC5891	omega-6 desaturase
21	147	6.1	382	Tl5042	omega-6 fatty acid
22	144	6.0	352	B69901	fatty-acid desatur
23	141.5	5.9	351	Tl1519	phosphatidylcholin
24	140	5.8	287	Tl1959	fatty-acid desatur
25	138	5.7	349	S43770	phosphatidylcholin
26	136	5.6	347	S43771	phosphatidylcholin
27	133	5.5	370	B83034	conserved hypothet
28	130	5.4	383	Tl0480	Delta12 fatty acid
29	123	5.1	385	T09880	omega-6 desaturase

ALIGNMENTS

RESULT 1

T26280

linoleoyl-CoA desaturase (EC 1.14.99.25) W08D2.4 - Caenorhabditis elegans  
N:Alternate names: Delta6 fatty acid desaturase  
C:Species: Caenorhabditis elegans

C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 21-Jul-2000  
C:Accession: T26280; T37238

R:Swinburne, J.; Ainscough, R.

submitted to the EMBL Data Library, March 1996

A:Reference number: Z20188

A:Accession: T26280

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-473 <WIL>

A:Cross-references: EMBL:Z70271; PIDN:CAA94233.1; GSPDB:GN00022; CESP:W08D2.4  
A:Experimental source: clone W08D2

R:Napier, J.A.; Hey, S.J.; Lacey, D.J.; Shewry, P.R.

Biochem. J. 330, 611-614, 1998

A:Title: Identification of a caenorhabditis elegans Delta6-fatty-acid-desaturase by h

A:Reference number: Z21637; M01D:98149727

A:Accession: T37238

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-38,69-430,'V',432-473 <NAP>

A:Cross-references: EMBL:AF031477; NID:g3088519; PIDN:AAC15586.1; PID:g3088520

C:Genetics:

A:Gene: CESP:W08D2.4

A:Map position: 4

A:Introns: 13/3; 39/2; 234/3; 277/3; 378/1; 413/3

C:Superfamily: Caenorhabditis elegans Delta6 fatty acid desaturase

C:Keywords: alternative splicing; oxidoreductase; unsaturated fatty acid biosynthesis

Query Match 98.8%; Score 2382; DB 1; Length 473;  
Best Local Similarity 93.4%; Pred. No. 1.3e-182;  
Matches 442; Conservative 1; Mismatches 0; Indels 30; Gaps 1;

Qy 1 MYVDKNASGLRMKVDGKNLYSEELVKKHGPGAVTEQ----- 37

Db 1 MYVDKNASGLRMKVDGKNLYSEELVKKHGPGAVTEQSYIPPLNKNIEIRGITTRGSSN 60

Qy 38 -----YRNSDATHIFHAFHEGSSQAYKOLDLLKKGHEDEFELEKLEKLDKVDINVS 90

Db 61 ALDILYFRNSDATHIFHAFHEGSSQAYKOLDLLKKGHEDEFELEKLEKLDKVDINVS 120

Qy 91 AYDVSAQEKKMKVESFEKLRKLDHDDGLMKANETVFLFKASTLSIMAFAYLQYLGWYI 150

Db 121 AYDVSAQEKKMKVESFEKLRKLDHDDGLMKANETVFLFKASTLSIMAFAYLQYLGWYI 180

Qy 151 TSACLLALAWQOFGWLTFEFCQQPTKNRPLNDTISLFFGNFQFSRDWKKDKINTHHA 210

Db 181 TSACLLALAWQOFGWLTFEFCQQPTKNRPLNDTISLFFGNFQFSRDWKKDKINTHHA 240



A;Residues: 1-458 <SPE>  
A;Cross-references: EMBL:X87143; NID:g1040728; PIDN:CAA60621.1; PID:g1040729  
C;Superfamily: cytochrome b5 core homology  
C;Keywords: heme; iron; metalloprotein; oxidoreductase  
F;16-90/Domain: cytochrome b5 core homology <CB5>  
F;51,74/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 19.1%; Score 460.5; DB 2; Length 458;  
Best Local Similarity 28.9%; Pred. No. 5.2e-29;  
Matches 131; Conservative 71; Mismatches 191; Indels 61; Gaps 16;

QY 6 NASGLRMKVDGKWLYLSEELVKKHGG-AVIEQYNRSDATHIFAHHEGSSOAYKOLDLL 64  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 29 NPNDLWISILGK-VYNVTWEAKHPPGDALINLAGODVDAFTAFHPGT--AWKHLDKL 85  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 65 KKHGEHDEFLEKLEKRDLKDVINSAIDVSVAQEKKWFESFEKLRLQKLDHGDKLMKANET 124  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 86 -----FTGYHLKDQVDSISRDKLAASEFAKAGMEKKGH 121  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 125 YFLFK-AISTLSIMAFYLOYLG-W-YITSACLLALAQCGWLTHFCHQOPTKNRP 180  
:: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 122 GVYSLCFVSLLSACVYGVLVSGSFWIHMLSGAILGLANMQIAYLGDHDAGHYQMATR 181  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 181 LNDTISLFFGNLQGFSRWKDKHNTHHAATNVIDHDGIDLAP-----LFASFPGD 233  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 182 WNKPAFIGINCITGISIAWKKTWNAAHHACNSLDYPDLOQLPMLAVSKLFNSTSV 241  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 234 LCYKASFEKAILKVIPYQHLYFTAMPLMRFWSQTGSQVOWFKEKOMEXKYVORNAFWE 293  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 242 FYGRQLTFDFLARFEVSYQHLYVPINCVARNVLYLTILLISKR----KIPDRG--- 293  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 294 QATIVG---HWAVFYOLFLLPTWPLRVAFYFIISOMGGGLLIAHV-VTFNHNSVDKY--P 347  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 294 -LNILGTLIEWFPPLLVSRPNWPVERVALVVSFCVTG--IQHIQFTLNFHSBGVVYP 350  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 348 ANSRILNFAALOILTTRNTTPSPFDNLWGGLNYQTEHHLFTPMPRCNLACVKKYVKW 407  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 351 PKG---DNWFEKOTRGTDITACSSMDWFGGLQGFQEHHLFPLPRCHLRKSIPICREL 407  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 408 CKENNPYLVDYDFDGVMNNLOQLKNMAEHTQAK 441  
|| : |||| : || : || : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 408 CKKNLPYVLSLFDVANVTTLKTLRTAA--LQAR 439  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 6  
TS0555  
delta-8 sphingolipid desaturase [imported] - rape  
C;Species: Brassica napus (rape)  
C;Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 15-Sep-2000  
C;Accession: T50555  
J.R.Sperling, P.; Zaehringner, U.; Heinz, B.  
J. Biol. Chem. 273, 28590-28596, 1998  
A;Title: A sphingolipid desaturase from higher plants. Identification of a new cytoch  
A;Reference number: Z22986; MUID:99003197  
A;Accession: T50555  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A;Residues: 1-449 <SPE>  
A;Cross-references: EMBL:AJ224160; PIDN:CAAI1857.1  
A;Experimental source: cultivar Drakkar  
C;Genetics:  
A;Gene: slcl

Query Match 18.4%; Score 443.5; DB 2; Length 449;  
Best Local Similarity 28.4%; Pred. No. 1.1e-27;  
Matches 127; Conservative 73; Mismatches 192; Indels 55; Gaps 16;

QY 10 LRMKVDGWLYLSEELVKKHGG-AVIEQYNRSDATHIFAHHEGSSOAYKOLDLKKHG 68  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 24 LWISIQQK-VYDVSHWVKSPGGEEAALLNLAGODVDAFTAFHPGT--AWRHLENL----76  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 69 EHDEFLEKLEKRLKVDINVSADVSAQEKMWSESEKLRQKLHDDGLMKANETPLF 128  
Db 77 -HNGYHVR-----DHHVS--DVS-----RDYRLAABFSKGLFDKKGHVTLY 116  
Qy 129 KAISTLSMAFAFY----LQVLGWYITSACLLALAWQOFGWLTDFCHQOQTKNRPLNDT 184  
Db 117 TLTCVAAMLAAYGVVACTSIWAHLISAVLLGLLWQISAYVGHSDSHYNTSTKPCNKL 176  
Qy 185 ISLFGNPLQFSRDWKKDKHNTHAATNVIDHDGIDDLAPLFA----FIPGDLCY--- 237  
Db 177 VOLLGNCITGISTAWKWTNAHHISCONSLDHPDLQHIPVLAVSNKFFKSWTSRFYGR 236  
Qy 238 KASFEKAILKIVPYOHLVFTAMPLMRLFSWTGOSVQVWFKENQMEYKVIYQARNAFWEQATI 297  
Db 237 KLTFDPLARFLISQHWFSYDPMCVGRINLFIQTLLLLFSRR-----YVPDRALNIAGI 290  
Qy 298 VGHNAWVYQFLPPLTWPLRVAYFIISOMGGGLLIAHV-VTFNHNVDKY--PANSRLIN 354  
Db 291 LVFWTFEPLVSLFNPWQERIIIFVLSM--AVTAIQHVQFCNLNHFADVYTGPPNG---N 345  
Qy 355 NFAALQILTRNNMTPSPFIDMWGLNGYQIEHHLFPMPRCNLNACVKYKWECKENNL 414  
Db 346 DWEKQTAGTLDISCRSYMDWFFGLQFQLEHHLFPRLPRCHLRGVSVPVQELCKKHNL 405  
Qy 415 YLVDDYFDGYAMNLOQLKNMAEHOAK 441  
Db 406 YRSLSWNEANYTLRLTKAA--VQAR 430

RESULT 7  
T47950  
delta-8 sphingolipid desaturase (EC 1.14.99.-) [validated] - Arabidopsis thaliana  
N:Alternate names: protein F2A19.180  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 02-Sep-2000  
C:Accession: T47950; T51848  
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quet  
submitted to the Protein Sequence Database, January 2000.  
A:Reference number: 224480  
A:Accession: T47950  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-449 <DEH>  
A:Cross-references: EMBL:AL132962; PIDN:CAB71088.1  
A:Experimental source: cultivar Columbia; BAC clone F2A19  
R:Sperling, P.; Zaehring, U.; Heinz, E.  
J. Biol. Chem. 273, 28580-28596, 1998  
A:Title: A sphingolipid desaturase from higher plants. Identification of a new cytochrome  
A:Reference number: 222986; MUID:99003197  
A:Accession: T51848  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-449 <SPE>  
A:Cross-references: EMBL:AJ224161; PIDN:CAAL1858.1  
A:Experimental source: cultivar Columbia; mainly green parts, some flowers, few roots  
C:Genetics:  
A:Gene: sls1  
A:Map position: 3  
A:Note: F2A19.180  
C:Function:  
A:Description: (EC 1.14.99.-); delta-8 sphingolipid desaturase [validated, MUID:99003197]  
C:Keywords: oxidoreductase; unsaturated fatty acid biosynthesis

Query Match 18.08; Score 433; DB 2; Length 449;  
Best Local Similarity 30.38; Pred. No. 7.9e-27;  
Matches 138; Conservative 61; Mismatches 183; Indels 74; Gaps 20;  
Qy 6 NASG-LRMKVQKWLYLESEELVKKHPGG-AVIEQYRNSDATHIFAFHEGSSQAYKQLDL 63  
Db 19 NKSGLMTAIOGK-VYNVSDWIKTHPGGDTVILNLVGQDVTDAFIAFPGT--AWHHLDH 75  
Qy 64 LKKHGEHDEFLEKLEKRLKVDINVSADVSAQEKMWSESEKLRQKLHDDGLMKANE 123

Db 76 LFT-GYHIR-----DFQVS--EVS-RDYRMAAEFRKL-----GL----- 106  
Qy 124 TYLEFKAISTLSIMAF--AFYLOVLGWYITSAC-----LLALAWQOFGWLTDFCH 170  
Db 107 --FENKGHVTLTYTAFVAAFMELGVL--YGVLACTSVFAHQIAAALGLLMTQSIYIGHDS 162  
Qy 171 CHQOQTKNRPLNDTISLFFGNFLQFSRDWKKDKHNTHAATNVIDHDGIDDLAPLFA-- 228  
Db 163 GHYVIMSNKSYNRRFAQLLSGNCLTGISTAWKWTNAHHHLACNSLDYDPLQHIPVFAVS 222  
Qy 229 --FIPGDLCY---KASFEKAILKIVPYOHLVFTAMPLMRLFSWTGOSVQVWFKENQMEY 283  
Db 223 TKFSSLTSTREYDRKLTDFPVARELVSYQHFYYPVCMCFGRINLFIQTLLLLFSKREVPD 282  
Qy 284 KVIYQARNAFWEQATIVGHNAWVYQFLPPLTWPLRVAVFIISOMGGGLLIAHV-VTFNHN 342  
Db 283 RAL-----NFAGILVFTWTFPLLVSLCPNWPFR--FFVFTSTVTALQHIQFTLNHFA 334  
Qy 343 VDKY---PANSRLINNFAALQILTRNNMTPSPFIDMWGLNGYQIEHHLFPMPRCNLNA 399  
Db 335 ADVVVGPTGS---DWEKQAGTIDISCRSYMDWFFGLQFQLEHHLFPRLPRCHLRK 390  
Qy 400 CVKYVKECKENNLPLYLVDDYFDGYAMNLOQLKNMA 435  
Db 391 VSPYVQELCKKHNLPRYSMSWFEANVLTLTKTAA 426

RESULT 8  
JC7556  
linoleoyl-CoA desaturase (EC 1.14.99.25) - Mucor rouxii  
N:Alternate names: delta6-desaturase  
C:Species: Mucor rouxii  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7556  
R:Laoteng, K.; Mannontarat, R.; Tanticharoen, M.; Cheevadhanarak, S.  
Biochem. Biophys. Res. Commun. 279, 17-22, 2000  
A:Title: Delta6-desaturase of Mucor rouxii with high similarity to plant delta6-desat  
A:Reference number: JC7556; MUID: 20563795  
A:Accession: JC7556  
A:Molecule type: DNA  
A:Residues: 1-523 <LAO>  
A:Cross-references: GB:AF290983  
A:Experimental source: strain ATCC 24905  
C:Comment: This enzyme, a membrane-bound key enzyme, is responsible for the transform  
C:Keywords: oxidoreductase; transformation

Query Match 17.5%; Score 421; DB 2; Length 523;  
Best Local Similarity 26.1%; Pred. No. 8.8e-26;  
Matches 127; Conservative 62; Mismatches 203; Indels 94; Gaps 13;  
Qy 19 LYLSEELVKKHPGG-AVIEQYRNSDATHIFAFH-----SSQAYKQ-LDLLKKHGEHDEFLEKLEKRLDKVDI 51  
Db 44 VYRVNFMKHPGGEAALRSALGRDVTDETRTMHPQVYKELINLYCIGDYMPDVIIRPAS 103  
Qy 52 -----EG---SSQAYKQ-LDLLKKHGEHDEFLEKLEKRLDKVDI 87  
Db 104 MKQOHTTKPKEDKPLVTATWEGGFTVQAYDDAIQDLKKHSHDLIKDAVLQKDLNGDOI 163  
Qy 88 NVSAYDSVAQEKMWSESEKLRQKLHDDGLMKANETVFLFKASTLSIMAFAYLQVLG 147  
Db 164 R-----NAYRKLEALEYAKGLFCNKYKAREGCRYTLTFLSLWFTLKG 208  
Qy 148 --W-YITSACLLALAWQOFGWLTDFCHQOQTKNRPLNDTISLFFGNFLQFSRDWKK 203  
Db 209 TETHYMAAGAAFMAMFWHLQVLTADAGHNEITCKSEIDHVIQVIGVILANFISGLSLGWMKD 268  
Qy 204 KHNTHAATNVIDHDGIDDLAPLFAF-----IPGDLCYKASFEKAILKIVPYQHLF 256  
Db 269 NHNVHHVTHNHPEDDPIQHVFPFAITTKFPNNIYTYKRVLPDAASFEFVRHQHLY 328  
Qy 257 TAMPLMRLFSWTGOSVQVWFKENQMEYKVIYQARNAFWEQATIVGHNAWVYQFLPPLTWPL 316

Db 329 YLILSFGFNLRLSLTCKNV-----RTRTLELVGITFFVFWFGSLSLSTLPTWNI 382  
Qy 317 RVAYFLISOMGGGLIAHV-VTFNH---NSVDKVPANSRLNNFAALOILTRNMTSPDF 372  
Db 393 RIAYIMVSYMTTFD--HVQITLSHFGMSTEDRGDEP-----FPKMLRTTMDVDCPEW 435  
Qy 373 IDLWGGUNYQIEHLLFPTMPRCNLNACVKYKWKCKENNLPLYLVDDYFDGYAMNLOQLK 432  
Db 436 HDWFHGGLOYQAVHLLFPLRHLRQCPLVKKFCDEVLGHYMYNFTSGNGVVLGTLK 495  
Qy 433 NNAEHI 438  
Db 496 SVADQV 501

RESULT 9  
Tl3155  
linoleoyl-CoA desaturase (EC 1.14.99.25) [validated] - human  
N:Alternate names: Delta6 fatty acid desaturase; protein DKFp586C201.1  
C:Species: Homo sapiens (man)  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jun-2000  
C:Accession: Tl3155; T08765  
R:Cho, H.P.; Nakamura, M.F.; Clarke, S.D.  
J. Biol. Chem. 274, 471-477, 1999  
A:Title: Cloning, expression, and nutritional regulation of the mammalian Delta-6 desaturase  
A:Reference number: 217612; MUID:99085046  
A:Accession: Tl3155  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-444 <CHO>  
A:Cross-references: EMBL:AF126799; NID:g4406527; PID:g4406528; PIDN:AAD20018.1  
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: 216471  
A:Accession: T08765  
A:Molecule type: mRNA  
A:Residues: 'RTRG', 138-428, 'D', 430, 'M', 432-444 <WAM>  
A:Cross-references: EMBL:AL050118  
A:Experimental source: adult uterus; clone DKFp586C201  
C:Genetics:  
A:Gene: GDB:FADSD6  
A:Cross-references: GDB:9956652  
A:Note: DKFp586C201.1  
C:Superfamily: cytochrome b5 core homology  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase; unsaturated fatty  
F:18-94/Domain: cytochrome b5 core homology <CB5>  
F:53,76/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 16.3%; Score 393.5; DB 2; Length 444;  
Best Local Similarity 27.8%; Pred. No. 1.1e-23;  
Matches 120; Conservative 61; Mismatches 166; Indels 85; Gaps 17;  
Qy 17 KWLYLSEELV-----KHGG-AVIEQYRNSDATHIFAFHFGSSQAYKOLDLLKKHGE 69  
Db 35 RNLVIDRKVYNTKWSIQHPGQGRVIGHYAGEDATDAFRAHP-----DL----- 79  
Qy 70 HDEFLEKLEKRLDKVDINVSAYDSVAQEKKMWESFEKLRKQKLDHDDGLMKRANETYFLFK 129  
Db 80 --EFVGKFL-RPLLIGELAPEEPSQDHCKSKITEDDFRALRKRTAEDMNLFKTNHVFLL 136  
Qy 130 AISTLSIMAFAYLQYL---GWYIT--SACLLALAWQOFGWLTFEFCQQPTKNRPLNDT 184  
Db 137 LAHIIALESIAFWTFYFGNGWIPTLITAFVLATSOAAGWLQHDYGHLSYVRKPKWNHL 196  
Qy 185 ISLFGFNLOGFSRDWKKDKHNTHAATNVIDHGDIDLAPLFAFIPGDLCKYKASFKA 244  
Db 197 VHKFVIGHLKASANWNNRHFQHHAKPNIFHKDPDVNN--LHVFVLGEW--QPIEYCKK 252  
Qy 245 ILKIVPY--QHLTYTAMLPMLRFSWVGOSVQWFEKQMEYKVVORNAFWEQATIVGH-- 300  
Db 253 KLKYLIPYHNOHEYFLIGPPLLIIPM-----YFOYQIIM-----TMIVHKN 292

Qy 301 -----WAWVFYQLFLLPTWPLRVAYFIISQMGGLLIAHVTF-----NHNSV 343  
Db 293 WYDLAWAVSYIRFTIYIPF---YCIL---GALLFLNFIRFLESHWFVWVQTNHIVM 345  
Qy 344 DKYPANSRLNNFAALQILTRNMTSPSPFDLWGLNGLYQIEHLLFPTMPRCNLNACVKY 403  
Db 346 E---IDQEAYRWFSQLTATCNVEQSFNDWFSGLNFQIEHLLFPTMPRNLHAKIAPL 402  
Qy 404 VKEMCKENNLPLY 415  
Db 403 VKSLCAKHGIEY 414  
RESULT 10  
JG0180  
Delta6 fatty acid desaturase (EC 1.14.99.-) [imported] - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: JG0180  
R:Aki, T.; Shimada, Y.; Inagaki, K.; Higashimoto, H.; Kawamoto, S.; Shigeta, S.; Ono, Biochem. Biophys. Res. Commun. 255, 575-579, 1999  
A:Title: Molecular cloning and functional characterization of rat delta-6 fatty acid  
A:Reference number: JG0180; MUID:99160394  
A:Accession: JG0180  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-444 <AKI>  
A:Cross-references: DBJ:AB021980; NID:g4514721; PIDN:BAA75496.1; PID:g4514722  
C:Superfamily: cytochrome b5 core homology  
C:Keywords: heme; iron; metalloprotein; oxidoreductase; unsaturated fatty acid biosyn  
F:18-94/Domain: cytochrome b5 core homology <CB5>  
F:53,76/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 16.2%; Score 389.5; DB 2; Length 444;  
Best Local Similarity 28.1%; Pred. No. 2.4e-23;  
Matches 123; Conservative 60; Mismatches 158; Indels 97; Gaps 19;

Qy 17 KWLYLSEELV-----KHGG-AVIEQYRNSDATHIFAFHFGSSQAYKOLDLLKKHGE 69  
Db 35 RNLVIDRKVYNTKWSQRHPGGRVIGHYSGEDATDAFRAH-----LDLDFVGK---- 84  
Qy 70 HDEFLEKQL-----EKRLDKVDINVSAYDSVAQEKKMWESFEKLRKQKLDHDDGLMKAN 122  
Db 85 --FLKPLLIGELAPEEPSLDR-----GKSSQITEDFRALKKTKTAEDMNLFKTN 129  
Qy 123 EYTFLEKAI STL SIMAFA--FYQLY--GWY--ITSACLLALAWQOFGWLTFEFCQQPT 176  
Db 130 HLFF-FLLSHIIIVMESIAFWFILSYFGNGWIPTVITAFVLATSOAAGWLQHDYHLSVY 188  
Qy 177 KNRPLNDITSLFGFNLOGFSRDWKKDKHNTHAATNVIDHGDIDLAPLFAFIPGDLCK 236  
Db 189 KKSINWNVHVKFVIGHLKASANWNNRHFQHHAKPNIFHKDPDI--KSLHVFVLGEW-- 244  
Qy 237 KYASEKAILKIVPY--QHLTYTAMLPMLRFSWVGOSVQWFEKQMEYKVVORNAFWEQ 294  
Db 245 QPLEYGGKKLKLPLYNHNOHEYFLIGPPLLIIP-----MYFOYQIIMTMRRRDWDL 296  
Qy 295 ATIVGHWAWVFYQL-----FLLPTWPLRVAYFIISQMGGLLIAHVVT 337  
Db 297 ANAISYARFFTYTIPFYGILGALVFLNFIRFLESHW-----FVWVTQMN-----HIVM 345  
Qy 338 FHNHSDVKYPANSRLNNFAALQILTRNMTSPFDLWGLNGLYQIEHLLFPTMPRCNL 397  
Db 346 --EIDLHY-----RDWFSQLAATCNVEQSFNDWFSGLNFQIEHLLFPTMPRNL 396  
Qy 398 NACVKYKWKCKENNLPLY 415  
Db 397 HKIAPLVKSLCAKHGIEY 414

RESULT 11







---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 01:31:09 ; Search time 2744.86 seconds  
(without alignments)  
11146.141 Million cell updates/sec

Title: us-09-555-093-1  
Perfect score: 1462  
Sequence: 1 gctccacaaatgtgtctg.....aataaacctgttttcaa 1462

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
-----							

1	1451	99.2	1463	3	AF031477	AF031477
2	1451	99.2	1463	6	AX003603	Caenorhab
3	832.8	57.0	30748	3	CEW0802	Sequence
4	350.4	24.0	1461	3	AF114440	Caenorhab
5	348.2	23.8	1344	3	AF078796	Caenorhab
6	348.2	23.8	1344	6	AX020906	Sequence
7	112.8	7.7	1275	8	AF139720	Euglena g
8	98.6	6.7	1374	8	AF465282	Mortierel
9	97	6.6	1374	6	AR080598	Sequence
10	97	6.6	1617	6	AR080598	Sequence
11	97	6.6	1617	6	AR098439	Sequence
12	97	6.6	1617	6	AR136018	Sequence
13	97	6.6	1617	6	AF110510	Mortierel
14	96.2	6.6	38060	3	CEP13F2	Caenorhabd
15	90.6	6.2	1374	8	AF307940	Mortierel
16	90.6	6.2	1374	8	AF465281	Mortierel
17	90.6	6.2	1947	8	AF465283	Mortierel
18	89	6.1	1590	8	AB020032	Mortierel
19	78.8	5.4	1546	8	AB052086	Mucor cir
20	76.2	5.2	2012	6	AX069275	Sequence
21	76.2	5.2	2012	8	PPAJ2980	Physcomit
22	75.4	5.2	1572	8	AF290983	Mucor rou
23	75.4	5.2	1572	8	AF296076	Mucor rou
24	68.8	4.7	1856	8	AF005096	Ricinus c
25	68.4	4.7	1071	8	AF307941	Mortierel
26	68.4	4.7	1644	8	AF307942	Mortierel
27	68.4	4.7	1684	6	AR076814	Sequence
28	68.4	4.7	1684	6	AR084177	Sequence
29	68.4	4.7	1685	6	AR020904	Sequence
30	68.4	4.7	1685	6	I38430	Sequence 4
31	68.4	4.7	1685	8	AF007561	Borago of
32	68	4.7	1687	8	BOU79010	Borago Offi
33	62.6	4.3	1478	8	AY055117	Echium qe
34	62	4.2	1467	6	AX058832	Sequence
35	62	4.2	2040	6	AX058830	Sequence
36	62	4.2	2040	8	CPU250734	Ceratodon
37	61.4	4.2	1450	8	AY055118	Echium pi
38	56.2	3.8	1678	6	AX007241	Sequence
39	56.2	3.8	1704	8	AF428420	Arabidops
40	56.2	3.8	1705	8	ATAJ4161	Arabidops
41	56.2	3.8	95993	8	ATF2A19	Arabidops
42	55.8	3.8	2160	6	AX058840	Sequence
43	55.8	3.8	2160	8	CPU250735	Ceratodon
44	51.8	3.5	1141	6	AX083744	Sequence
45	51.8	3.5	1508	10	AF126798	Mus muscu

ALIGNMENTS

RESULT	1
AF031477	
LOCUS	AF031477 1463 bp mRNA linear INV 02-MAY-1998
DEFINITION	Caenorhabditis elegans delta6-fatty-acid-desaturase mRNA, complete cds.
ACCESSION	AF031477
VERSION	AF031477.1 GI:3088519
KEYWORDS	
SOURCE	Caenorhabditis elegans.
ORGANISM	Caenorhabditis elegans
REFERENCE	1 (bases 1 to 1463)
AUTHORS	Savanova, O., Smith, M.A., Lapinskas, P., Stobart, A.K., Dobson, G., Christie, W.W., Shewry, P.R. and Napier, J.A.
TITLE	Expression of a borage desaturase cDNA containing an N-terminal cytochrome b5 domain results in the accumulation of high levels of delta6-desaturated fatty acids in transgenic tobacco
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 94 (8), 4211-4216 (1997)
MEDLINE	97268723
REFERENCE	2 (bases 1 to 1463)
AUTHORS	Napier, J.A., Hey, S.J., Lacey, D.J. and Shewry, P.R.
TITLE	Identification of a caenorhabditis elegans

April 97



ORGANISM

Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 1463)

REFERENCE

Napier,J.A.

AUTHORS

Desaturase genes and their use

TITLE

Patent: WO 99/27111-A 1 03-JUN-1999;

JOURNAL

UNIV BRISTOL (GB); NAPIER JOHNATHAN A (GB)

FEATURES

Location/Qualifiers

source

1..1463

/organism="Caenorhabditis elegans"

/db\_xref="taxon:6239"

11..1342

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC04885.1"

/db\_xref="GI:9927414"

/translation="MVVDKNASGLRMKVDGKWLILSEELVKKHGGVIEQVNRSDAT  
HIFHAFHSGSQAQYKQDLKKHGHDEFLEKLEKLDKVDINVSADVSVAQEKKM  
VESFEKLRKQDLHDKANETFLKALSTLSIMAFYLYQLGYITSACLALAW  
QQFGLTHFCHQDPKRNPLNDTISLFPNPLQGFSDWMDKDKHNTTHAATNVIDHD  
GDIDLAFATIPGLCKYKASFKAIIKIVPYOHLXFTAMLPLMRFESWTSQVOWVF  
KENOMEYKVYORNAFWEGATLVGHWAIFYQLFLPLTWPLRVAYFIIISOMGGGLIAH  
VYTNHNSVDKIPANSRLNNFALQILITRNMTSPFIDMLWGLNTOIEHHLPPTM  
PRCLNACVKYKWKCKENNLPLYLVDDYFDGYAMNLOQLKMAEHIOAKAA"

BASE COUNT 445 a 288 c 277 g 453 t

ORIGIN

Query Match

Best Local Similarity 99.2%; Score 1451; DB 6; Length 1463;

Matches 1462; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 gctcaccacaaatgctgcgcacaaagaatcctccgggttcgaatgaagtcgatgca 60  
DB 1 GCTCACCACAAATGGTCTCGACGAAGATCCCTCCGGGCTTCGAATGAAGTCGATGGCAA 60  
QY 61 atggtctaccttagcgagggaattggtgaagaaacatccaggaggagctgttattgaaca 120  
DB 61 ATGGGCTCTACCTTAGCGAGGAATTGGTGAAGAAACATCCAGGAGGAGCTGTTATTGAACA 120  
QY 121 atatagaattcggatgctactcatattttccacgcttccacgaggatcttccagc 180  
DB 121 ATATAGAAATTCGGATGCTACTCATATTTTCCACGCTTTCCACGAGGATCTTCTCAGGC 180  
QY 181 ttataagcaacttgacctcttgaaagacacgagagacacgatgaattccttgagaacaa 240  
DB 181 TTATAGCAACTTGACCTTCTGAAAAGCAGCGAGACGACGATGAATTCCTTGAGAAACA 240  
QY 241 attggaagagagacttgacaaagtgtgatcaatgtatcagcatatgatgtcagttgttc 300  
DB 241 ATTGMAAAGAGACTTGACAAAGTTGATATCAATGTATCAGCATATGATGTGAGTTCG 300  
QY 301 acaagaagaagaatggttgatcattccgaagaatacagacgaagcttcgatgatg 360  
DB 301 ACAAGAAAGAAATGGTTGAATCATTCGAAAACATACGACAGAAGCTTCATGATGATGG 360  
QY 361 attaatgaagcaaatgaacatatcttctgtttaaagcgatttcaaaccttcaattat 420  
DB 361 ATTAATGAAGCAATGAACATATTTCTCTGTTAAGCGATTTCACACTTTCATATTAT 420  
QY 421 ggcatttgcatcttctcagtatcttgatgggttatattactcttgcatgtttattagc 480  
DB 421 GGCAATTTGCATTTTATCTTCAGTATCTTGGATGGTATATTACTTCTGCATGTTATTAGC 480  
QY 481 acttgcatggaacaattcggatggttaacacatgatgtctctgccatcaacagcccaaaa 540  
DB 481 ACTTGCAATGGCAACAATTCGGATGTTTAACACATGATGTTCTGCCATCAACAGCCAAACA 540  
QY 541 gaacagaccttgatgatactattcttcttcttcttcttcttcttcttcttcttcttctt 600  
DB 541 GAACAGACCTTGAATGATGATGATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 600

QY 601 aagagattggtgaagcaagcatacaactcatcagctgcacaaatgtaattgatca 660  
DB 601 AAGAGATTGGTGAAGCAAGCAAGCATAACTCATCAGCTGCCACAAATGTAATTGATCA 660  
QY 661 tgacggtgatatacgacttggcaccacttttcgattatttccaggagatttgcgaagta 720  
DB 661 TGACGGTGATATCGACTTGGCACCACCTTTTCGCATTTTATCCAGGAGATTGTCGAAGTA 720  
QY 721 taagggcagctttgaaaaagcaattctcaagattgtaccatatcaacatctctatttcc 780  
DB 721 TAAGGCCAGCTTTGAAAAGCAAAATCTCAAGATTGTACCATATCAACATCTCTATTTCAC 780  
QY 781 cgcaatgcttccaatgctccgtttctctcatggactggcagtgctcagttcaatgggtattcaa 840  
DB 781 CGCAATGCTTCCAATGCTCCGTTTCTCATGGACTGTGCTCAGTCAGTTCAATGGGTATTCAA 840  
QY 841 agagaatacaaaatggagatacaaggctctatcaaaagaatgaattctgggagcaagaacaat 900  
DB 841 AGAGAAATCAAAATGGAGTACAAAGTCTATCAAGAAATGCAATCTGGGAGCAAGCAAAAT 900  
QY 901 tgttgacattgggcttggtggtattcttatcaattgttcttattaccacaatggccacttcg 960  
DB 901 TGTGACATTTGGGCTTGGGTATTCTATCAATTTGTTCTTTATTACCAACATGGCCACTTCG 960  
QY 961 ggttgcttattcttatttatttcaaaaatgggagagggccttttgattgctcacgtagtcac 1020  
DB 961 GGTGCTTATTCTTATTATTTCACAAATGGGAGGAGGCTTTTGTGATTGCTCAGCTAGTCAC 1020  
QY 1021 tttaaccataactctgttgataagtagtatccagccaattctcgaatttttaaacacttcgc 1080  
DB 1021 TTTCAACCACTAATCTGTTGATAAGTATCCAGCAATTTCTCGAATTTTAAACCAACTTCGC 1080  
QY 1081 cgtctctcaaaatttgaccacacgacacatgactccatctccattcattgattggtgttg 1140  
DB 1081 CGCTCTTCAAAATTTGACCACACGCAACATGACTCCATCTCCATTCTTCAATTGATTTGGCTTG 1140  
QY 1141 gggtagactcaattatcagatcgagcaccacttgggtcccaacaatgccagcttgcgaatct 1200  
DB 1141 GGGTGGACTCAATTATCAGATCGAGCACCACCTTGTGTTCCCAACAATGCCAGTTGCAATCT 1200  
QY 1201 gaatgcttcggtgaataatgtgaagaatgggtgaagaagaataatctccttaccctcgt 1260  
DB 1201 GAATGCTTCGGTGAATAATGTGAAGAATGGTGCAAGAGAAATAATCTTCTTACCTCGT 1260  
QY 1261 cgatgactacttgacggtatgcaatgaatttgcaacaattgaaataatgctgctgagca 1320  
DB 1261 CGATGACTACTTTGACGGATATGCAATGAATTTGCAACCAATTTGCAAAATATGGCTGAGCA 1320  
QY 1321 cattcaagctaaagctgcctaaacaactctgggtgttccaaaagtgtttttcttg-tttttt 1379  
DB 1321 CATTCGAAGCTAAAGCTGCCTAAACAATCTGGGTGTTTCAAAAAGTTTTTCTGTTTTTTT 1380  
QY 1380 aaatttaattcttgaaattattgtttccgtcattcttctccattctccattcttcttctg 1439  
DB 1381 AAATTTAAATCTTGAATTAATTTGTTTTCGCTCATTTCTCTCCATTCCTCTCTCTCTG 1440  
QY 1440 agaaaataaaacctgtttttcaa 1462  
DB 1441 AGAATAAAACCTTGTTTTTCAA 1463

RESULT 3

CEW08D2

LOCUS

Caenorhabditis elegans cosmid W08D2, complete sequence.

DEFINITION

270271.1 GI:1236102

ACCESSION

HTG; Collagen; Human myeloid cell line protein like; Transporter

KEYWORDS

ATPase like; Tyrosine-protein kinase; WNT protein.

SOURCE

Caenorhabditis elegans.

ORGANISM

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE AUTHORS TITLE	1 (sites) none.				4202. .4479,4668. .5004,6413. .6748,6795. .6896,7376. .7420))
JOURNAL MEDLINE	Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium Science 282 (5396), 2012-2018 (1998)				/gene="w08d2.3"
REMARK	99069613				/note="predicted using Genefinder"
REFERENCE AUTHORS TITLE	2 (bases 1 to 30748) The <i>C. elegans</i> Sequencing Consortium. Swinburne,J. and Ainscough,R. Direct Submission Submitted (23-MAR-1996) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jes@sanger.ac.uk or rwnematode.wustl.edu Coding sequences below are predicted from computer analysis, using predictions from Genefinder (P. Green, U. Washington), and other available information. Current sequence finishing criteria for the <i>C. elegans</i> genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note. IMPORTANT: This sequence is not the entire insert of clone w08d2. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. The true left end of clone w08d2 is at 1 in this sequence. The true right end of clone w08d2 is at 5560 in sequence w07284. The true left end of clone w07f5 is at 30645 in this sequence. The start of this sequence (1. .108) overlaps with the end of sequence z81122. The end of this sequence (30645. .30748) overlaps with the start of sequence w07284. 960307: Dual-authors: June Swinburne & Rachael Ainscough For a graphical representation of this sequence and its analysis see: <a href="http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=w08d2">http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=w08d2</a> IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. Location/Qualifiers 1. .30748 /organism="Caenorhabditis elegans" /db_xref="taxon:6239" /chromosome="IV" /clone="w08d2" join(21. .59,907. .983,1289. .1784,2096. .2224,2272. .2572,2623. .2729,2775. .2957) /gene="w08d2.4" join(21. .59,907. .983,1289. .1784,2096. .2224,2272. .2572,2623. .2729,2775. .2957) /gene="w08d2.4"				/codon_start=1 /protein_id="CAA94232.1" /db_xref="GI:3880556" /translation="MPIEVNDFDPAYHPKPIFDRIVEAKETASKHPSLFITIHLMWYS LIGTWTFFISLSYSHSACYMISSAHLNLFNFAFNGLSGLVAVINFFVFFKQOQFLL ITCLILISACFLPAAHAHITLLCIVIFVFGTVGLISLAVINFFVFFKQOQFLL OFLHLGAFGVTAVMPYQSAPESEHEDCKSRNPVLPPTGPHIPSLKFKYLNPSLRD QQDSVSLKVLVCSVASIANIMTFFPFLIQAQSAFIYOLFVLSKQKQKQFYK GYFFQIFFMGRAVSLFFSDYMAPYFLVGSIFGCGCCCAISEQFLVLSGLIFGF LAPLLPALVIYNHLSLKSNTFTLLSGHSGHLIFPFLTLTNSNSRFLIAFNFV AIVTLVILVLSLHAQIERMOASQGLNLFVSRFLSGESLLKTRISRLPSLRP NSYRFRANRRGLNTRSPSPAPVPHSTEVSPDELSRLNEKPEIRSGSMLMTRY" complement(join(10141. .10236,10563. .10663,10723. .10849, 10969. .11112,11159. .11329,11379. .11518,11918. .12000, 12059. .12138,12192. .12324,12373. .12479)) /gene="w08d2.1" complement(join(10141. .10236,10563. .10663,10723. .10849, 10969. .11112,11159. .11329,11379. .11518,11918. .12000, 12059. .12138,12192. .12324,12373. .12479)) /gene="w08d2.1" /note="Similarity to Mouse WNTB protein (SW:WNTB_MOUSE), contains similarity to Pfam domain: PF00110 (wnt family of developmental signaling proteins), Score=366.8, E-value=1.6e-133, N=1" /codon_start=1 /protein_id="CAB61041.1" /db_xref="GI:6434322" /db_xref="SPTREMBL:Q9TVJ1" /translation="MORFICLIFLLVGVQSPSATYSTQOVVNWMLTLAFVGSN DYLPRPHYSTDREHFKELCRRLDGLNPQALCAENPFSPFVARGVREAIKCEK FKFERNCSSDDEVETRHGFKQDLILKRSANKRAFLNAIMAAISLVHSITGKNT GNLTCCGDSKSPQMRYQASDPSMRDQFSWGGSDNVPHGIRYAKKFLDOWETAQF DKTNDVAHVLRNNKTVGREAIQAQIRQCRCGHVSGCEFTCLQMKFSQVSDLL KKRYDFAVQVTRKATRLRKRTERRKILPLRGEMAYVHRSPSYCEKNLTAGILGTS GRECIHNSYSSCDLCCGRGYNTRLEIROTCCKFVWCCEVKCKTCTEEVAVHTC K" Complement(join(14668. .14800,14856. .15035,15765. .16123, 16170. .16381,16426. .16813,16861. .17023,17068. .17261, 17308. .17638,17687. .17907,17954. .18057,18115. .18256, 18304. .18408,20054. .20173,20225. .20353,20399. .20477, 20536. .20636,20684. .21027,21092. .21143,21187. .21317, 21440. .21505)) /gene="w08d2.5" complement(join(14668. .14800,14856. .15035,15765. .16123, 16170. .16381,16426. .16813,16861. .17023,17068. .17261, 17308. .17638,17687. .17907,17954. .18057,18115. .18256, 18304. .18408,20054. .20173,20225. .20353,20399. .20477, 20536. .20636,20684. .21027,21092. .21143,21187. .21317, 21440. .21505)) /gene="w08d2.5" /note="Similarity to Yeast E1-E2 ATPase (SW:YED1_YEAST), contains similarity to Pfam domain: PF00122 (E1-E2 ATPase), Score=102.4, E-value=2.7e-28, N=4 CDNA EST Yk5f9.5 comes from this gene CDNA EST Yk10d12.5 comes from this gene CDNA EST Yk5f9.3 comes from this gene CDNA EST Yk10d12.3 comes from this gene CDNA EST Yk40h11.5 comes from this gene CDNA EST Yk131q11.3 comes from this gene CDNA EST Yk131q11.5 comes from this gene CDNA EST Yk133d6.5 comes from this gene CDNA EST Yk318f2.3 comes from this gene CDNA EST Yk318f2.5 comes from this gene CDNA EST Yk248h3.5 comes from this gene CDNA EST Yk260b8.5 comes from this gene CDNA EST Yk260b8.5 comes from this gene CDNA EST Yk286h5.5 comes from this gene CDNA EST Yk356g1.5 comes from this gene"
FEATURES source	gene	gene			
	CDS	CDS			
gene					
CDS					

```
/codon_start=1
/db_xref="GI:3880560"
/db_xref="SWISS-PROT:Q27533"
/translation="MTLESQDHTLTFAYRCPPTILFYALTVLTLGIFRLILHWKQ
KWDKMRVPCFEEAEYIIDHNHSELQPVLRKSNATIPTEGENMKRYPPELRWFY
YRLEFYWIDNLSDESDEISDRNCWKTSFEIANRIPCRSLNVAESNFGLLTSEI
SRLEFYGRNIVQLQRPILYLLVMEVITPFYVFQIFSVTVMYDEVAYIASLVILS
LGSIVMDVQIRTEQIRLRSVHSTSEVEIREGTEMTIGSDQVPGDILLIPHGCL
MQDSVLNMGVIVINESVLTCESVPIPKVALTDETNDSVFNIENKSNVLCGQTVLQ
TRFYRGKVAIVLRTAYSTLKQLVRSIMPKVPVDFRTKDLFKFLFLACISCGGF
IYTIIMRNVTLRRIIVRSLDIITIVPALPAAMSVGIINAQLRUKKEIFCISP
STINCAINVCEDKGTJTEDEGLDFHVRPVMASVAVNOEIQKLEKSNRTEPMGE
TELSRFGPDGLVLAIAATCHSLRINGVLHGDPLDLILFQKGTWMEBIGEDIE
EETQRFNVQSIILKPDKAESYVIRQFTFSSLRQMSVIFDPREDRDNMMLYS
APDQSLVELOQIDIVTAMCGDGAANDCAALKAHAGISLSDAEASTAAPFTSKVPD
IRCVPTISLGRAALVTSFGIFKYMAGYSLTOFVTVMHLYWISNILLTDQOFMYIDMFL
ITMFAALLGNTPAFYRLAHTPPTRLLSIASMTSVGOLIIIGVVQFIVFPTSQOPW
FTPYQPVDDDEVDKRSQGTALFCVSMFOYIILALVYSKGPFRGNLWSNKPYYKKK
RSIEAIDIVFTNSDHRPSINGVTSSTESTLSSAEGQQLHMTTSKNGKGGENPH
SSALFERLISIRIGEPWLTNPPIPHSLSEPEPEKLETTY"
join(18853..19023,19072..19859,19907..20024)
/gene="w0802.6"
/join(18853..19023,19072..19859,19907..20024)
/gene="w0802.6"
/note="predicted using GeneFinder
contains similarity to Pfam domain: PF01391 (Collagen
triple helix repeat (20 copies)), Score=73.1,
E-value=1.9e-18, N=2; PF01484 (Nematode cuticle collagen
N-terminal domain), Score=67.8, E-value=7.3e-17, N=1
CDNA EST yk308e7.3 comes from this gene
CDNA EST yk308e7.5 comes from this gene
CDNA EST yk385a8.3 comes from this gene
CDNA EST yk385a8.5 comes from this gene"
/codon_start=1
/protein_id="CAA94234.1"
/db_xref="GI:3880558"
/db_xref="SPTREMBL:Q23222"
/translation="MGIFSEFLYRMQVSSVTFVASSISAIALISCLLSSVYRNVD
YWTLEDDQMEVREKDLDDMWREIMHIQISLGSVFRARRPARGAYDEGDYKEADPPPTS
```

```
Query Match 57.0%; Score 832.8; DB 3; Length 30748;
Best Local Similarity 74.6%; Pred. No. 7.5e-174;
Matches 1339; Conservative 0; Mismatches 2; Indels 455; Gaps 5:

Qy 122 tatagaaattcggatgctactcaattttccacgctttccacgaaggatcttccaggct 181
|
|
|
Db 1284 TATAGAAATTCGGATGCTACTATATTTCCACGCTTTCCACGAAGGATCTTCTCAGGCT 1343
|
|
|
Qy 182 tataagcaacttgacctctgaaagacagagagagcagatgaattccttgagaacaa 241
|
|
|
Db 1344 TATAAGCAACTTGACCTCTGAAAAGACGCGAGAGCAGATGAATTCCTTGAGAAACAA 1403
|
|
|
Qy 242 ttggaagagacttgacaaagtgtatcaatgtatcagcatatgatgtcagttgttaca 301
|
|
|
Db 1404 TTGGAAGAGACTTGACAAAGTGTGATATCAATGTATCAGCATATGATGTCTAGTGTGCA 1463
|
|
|
Qy 302 caagaaagaaaaatggttgaaatcattcgaaaaactcacgacagaagcttcattgatgga 361
|
|
|
Db 1464 CAAGAAAGAAAAATGGTTGAATCATTCGAAAAACTACGACAGAAGCTTCATGATGGA 1523
|
|
|
Qy 362 ttaatgaagcaaatgaacatatcttcctgtttaagcgatttcacactttcaattatg 421
|
|
|
Db 1524 TTAATGAAGCAAAATGAACATATTTCTGTGTTAAAGCAATTTCAACACTTTTCAATTATG 1583
|
|
|
Qy 422 gcatttgcatatttcttcagtatcttgatgtgtattacttctgcattgtttattagca 481
|
|
|
Db 1584 GCATTGTCATTTATCTTCAGTATCTTGATGGTATATTTACTTCTGTCATGCTTTATTAGCA 1643
|
|
|
```

```
Qy 482 cttgatggcaacaattcggatggttaacacatgattctgccatcaacagcacaacaaag 541
|
|
|
Db 1644 CTTGCATGCCAACAAATTCGGATGGTTAACACATGAGTCTGCCATCAACAGCCACAAG 1703
|
|
|
Qy 542 aaacagaccttgaatgatatactattttcttctgttcttctgttaattcttacaagatttca 601
|
|
|
Db 1704 AACAGACCCTTGAATGATATACTATTTCTTGTCTTCTTGTGTAATTTCTTACAAGGATTTTCA 1763
|
|
|
Qy 602 agagatttgtgaaagaca- 620
|
|
|
Db 1764 AGAGATTGGTGAAGGACAAAGGTGAGAAATATGAGAGACTTGCAATACCGTATCCGTTTTC 1823
|
|
|
Qy 621 - 620
|
|
|
Db 1824 AAAGCTTTTCTGCTGCTGCTGACITCTCTTGAAAAAATAGGTGATAAAGGTATCATTG 1883
|
|
|
Qy 621 - 620
|
|
|
Db 1884 TGGGATTATTTTCTAAACTGTTTTTTTCCGTTAGACACAAAGCACAATAATATATA 1943
|
|
|
Qy 621 - 620
|
|
|
Db 1944 TTGATCATAGTGGTCTCAAGACCGTAACCCCTCTATTGACATAAGAAATCAITTCATTT 2003
|
|
|
Qy 621 - 620
|
|
|
Db 2004 CAATTTCTGATGATTATTTAAAAAAATTTCTTTTAAATAATTGGAATAAATAAAAA 2063
|
|
|
Qy 621 - 620
|
|
|
Db 2064 TGTTTATCCAGTCTCTTTAATATCAATTTACAGCATAACTCATCATCGCTGCCACAAATG 2123
|
|
|
Qy 651 taatgatcatcacggtgatatacgcattgacacacatttccatttattccagagatt 710
|
|
|
Db 2124 TAATGATCATCAGCGTGATGCTGACCTTGACACCACTTTTCGATTTATTTCAGGAGATTT 2183
|
|
|
Qy 711 tgtcaagtataaggccagctttgaaaagcaattctc- 748
|
|
|
Db 2184 TGTGCAAGTATAAGGCCAGCTTTGAAAAAGCAATTTCTCAAGGTAAAACTGAATTTTCTAT 2243
|
|
|
Qy 749 - 748
|
|
|
Db 2244 AACTTCTCTAATATATTAATTAATTTTAAAGATGTACCATATCAACATCTCTATTTTCA 2303
|
|
|
Qy 784 aatgcttccaatgctcgtttctcatgagtcagtcagtcagtcagtcagtcagtcagtcagtc 843
|
|
|
Db 2304 AATGCTTCCAATGCTCCGCTTCTCATGGACTGGTCACTGAGTCAATGGGTATTCAGAA 2363
|
|
|
Qy 844 gaatcaaatggagtagcaaggtctatacaagaatgcatctctgggagcaagcaaatgt 903
|
|
|
Db 2364 GAATCAAAATGGAGTACAAGGTCTATCAAGAAATGCAATTTCTGGAGCAAGCAAAATGCT 2423
|
|
|
Qy 904 tggcaattgggttgggttattctatacaattgttcttattaccacaatggccacttcgggt 963
|
|
|
Db 2424 TGGACATGGGTGGGTGATTTCTATCAATTTGTTCTTATTACCAACATGGCCACTTCGGGT 2483
|
|
|
Qy 964 tgcatttcttatttctcaaatgggagagcccttttgattgctcagtcagtcagtcagtcagtc 1023
|
|
|
Db 2484 TGCATTATTTTCAATATTTTCAAAATGGGAGAGAGCCCTTTTGTGTTGCTCAGCTAGTCACTTT 2543
|
|
|
Qy 1024 caaccataactgttgtataagatc- 1049
|
|
|
Db 2544 CAACCATAACTCTGTTGTAAGTATCCAGGTGAGATTTGACCGAAAGATGACTTTTTTTTT 2603
|
|
|
Qy 1050 - 1050
|
|
|
Db 2604 ACAAAAGACGAGTTCAGCCAAATTCGCAATTTTAAACAACCTTCGCCGCTCTTCAAAAT 2663
|
|
|
Qy 1094 ttgaccacacgacaataactcctcattcattgattgcttgggttgggtgagactcaat 1153
|
|
|
Db 2664 TTGACCACACGCAACATGACTCCATCTCCATTGATTGGCTTTGGGTTGGGTCGACTCAAT 2723
|
|
|
Qy 1154 ta- 1168
|
|
|
```







Db 991 TTCTGCTCTCTCATGTAGTTACTTTTCAATCAATTATTAGTGAGAGAGTTTGCATTGAGC 1050  
QY 1058 tctcgaaatttaaacacattccgcctcttcaaaattttgaccacacgaacatgactcca 1117  
Db 1051 TCGAACATCATGTCAAATATACGCTTCTCAAAATCATGACCACAAAGAAATATGAGACCT 1110  
QY 1118 tctccattcattgattggcttgggtggagctcaattatcagatcgagcaccacttgctc 1177  
Db 1111 GGAAGATTATTGACTGGCTTTGGGGAGGCTTAACATATCAGATGAGCACCATCTTTTC 1170  
QY 1178 ccaacaatgccagcttgcaatgaatgcttgcgtgaaatatgtgaaagaatgggtgcaa 1237  
Db 1171 CCAACGATGCCAGCACAACTTGAACACTGTATGCCACTTGTAAAGGAGTTTGAGCA 1230  
QY 1238 gagaataatctccttaccctcgatgactacttgacgggataatgcaatgaatttgcaa 1297  
Db 1231 GCAATGGTTTACCATACATGTTGGTCGACGATATTTCACAGGATTCGGCTTGAATTTGAG 1290  
QY 1298 caattgaaaaatatggc 1314  
Db 1291 CAATTCGGAATATTGC 1307

RESULT 6  
AX020906  
LOCUS AX020906 1344 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 2 from Patent WO9933958.  
ACCESSION AX020906  
VERSION AX020906.1 GI:10044591  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unclassified.  
REFERENCE 1. (bases 1 to 1344)  
AUTHORS Michaelson, L., Stobart, K. and Napier, J. A.  
TITLE Desaturase  
JOURNAL Patent: WO 9933958-A 2 08-JUL-1999;  
MICHAELSON LOUISE (GB); STOBART KEITH (GB); UNIV BRISTOL (GB);  
NAPIER JOHNATHAN A (GB)  
FEATURES  
source Location/Qualifiers  
1..1344  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 396 a 269 c 261 g 418 t  
ORIGIN

Query Match 23.8%; Score 348.2; DB 6; Length 1344;  
Best Local Similarity 55.8%; Pred. No. 9.7e-67;  
Matches 712; Conservative 0; Mismatches 553; Indels 12; Gaps 2;

QY 44 atgaaggtcgatgcaaatggctctaccttagcagggaattggtgaagaacatccagga 103  
Db 37 ATTAATAATTGATGGAAATATGGTGTCAAAATGACGATGCTCCTGAGATCAGATCCAGGT 96  
QY 104 ggagctgttatgaacaatatagaattcgatcgactcatatttccacgctttccac 163  
Db 97 GGTAGTGAATACTACTACTATAAAATATGGATGCCACTACCGTATTCCACACATTCAT 156  
QY 164 gaaggtctctcaggtcttaagaacaacttgacctctctgaaagaagcaagg-----agag 217  
Db 157 ACTGGTCTTAAAGAGCGGTATCAATGGCTGACAGAAATTGAAAAAGAGTGCCTACACAA 216  
QY 218 cagcatgaattccttgagaaacaattggaaagagacttgacaagtgtgatcaatgta 277  
Db 217 GAACGAGATCCAGATATTAAAGATGACCAATCAAGGAAATGATGATGTGAACATG 276  
QY 278 tcagcatgatgtcagttgtgcacaaagaaagaaatggttgaatcattcgcgaaaaacta 337  
Db 277 GGAACCTTCAATATTTCTGAGAAACGATCTGCCCAATAAATAAAGTTTCACTGATCTA 336  
QY 338 cgacagaagcttcatgatgatgattaatgaaagcaaatgaaacatatatttctgtttaa 397

Db 337 CGTATGCGAGTTTCGTGCAGAAGGACTTATGGATGGATCTCCTTTGTTCTACATTAGAAAA 396  
QY 398 gsgatttcaaacactttcaattatgcatctgatttatttatttcttcagttatcttgatggtat 457  
Db 397 ATTCCTTGAACAACATCTTCAAAATCTTTTGGCATCTTACCTTCAATACCACACATATTAT 456  
QY 458 attactctcgactgtttattagcacttgatggaacaattcggtatggttgaacacatgag 517  
Db 457 CTTCATCATGAGCTATTCTAATGGGAGTTGCGTGGCAACAATTTGGGATGGTTAATCCATGAA 516  
QY 518 ttctgcatcaacagacccaacaagaacacctttgaaatgatactatttctttgttcttt 577  
Db 517 TTGCGCATCATCAGTTGTTCAAAAACAGATACTACAATGATTTGGCCAGCTATTTCGTT 576  
QY 578 ggttaatttctcaaggaattttcaagagattggtggaagacaagaacataaacactcatcac 637  
Db 577 GGAACCTTTTTCAAAGGATTCATCTGTTGGTGGTGGAAAGAGCAGCACAATGTGCATCAC 636  
QY 638 gctgcccacaaatgtaattgatcatgacggtgatctatcgacttggcaccacttttcgcattt 697  
Db 637 GCAGCCACAATGTTGTTGGACGAGACGGAGATCTTGATTTAGTCCCATTCATGCTACA 696  
QY 698 atccaggagatttgcgaagtataaggccagctttgaaaaagcaattctcaagattgta 757  
Db 697 GTGGCAAGAACTCTCAACAATTTAT-----TCTCAGGATTCATGGGTTATGACTCTATTTC 750  
QY 758 ccatatcaacatctctatttccgcgaatgcttccaatgctccggtttctcatggactggt 817  
Db 751 AGATGGCAACATGTTCAATTTGGACATTCATTTACCATTCCCTCCGCTCTCTCGTGGCTCTT 810  
QY 818 cagtcagttccaatgggtatttcaaaagagaatcaaaaggagtaacaaggtctctatacaagaaat 877  
Db 811 CAGTCAATCATTTTGTAGTCAGATGCCAATCATATTATTATGACTATTACAGAAATACT 870  
QY 878 gcattctgggagcaagaacaatttggacattgggttggttggttattctatcaattgttc 937  
Db 871 GCGATTTATGAACAGGTTGGTCTCTCTTTGCACATGGGCTTGGTCAATTTGGTCAATTTGAT 930  
QY 938 ttattaccaaacatggccacttcggttgccttatttcttatttcttatttcaacaatggagagac 997  
Db 931 TTCTTACCCGATTTGGTCAACTAGAAATAATGTTCTTCTTCTTCTTCTCTCTTCTCTCTTCT 990  
QY 998 ctttgattgctcaogtagtctcactttcaaccataactctgttgaataagttatccagccaat 1057  
Db 991 TTCTCTGCTCTCATGCTAGTTACTTTCAATCATTTATTCAGTGGAGAAAGTTTGCATTGAGC 1050  
QY 1058 tctcgaattttaaacaaacttcgcgcctcttcaaaattttgaccacacgaacatgactcca 1117  
Db 1051 TCGAACATCATGTCAAATATACGCTTGTCTTCAAAATCATGACCACAAAGAAATATGAGACCT 1110  
QY 1118 tctccattcattgattggcttgggtggagctcaattatcagatcgagcaccacttgctc 1177  
Db 1111 GGAAGATTATTGACTGGCTTTGGGGAGGCTTAACTATCAGATGAGCACCATCTTTTC 1170  
QY 1178 ccaacaatgccagcttgcaatgcttgcgtgaaatatgtgaaagaatgggtgcaa 1237  
Db 1171 CCAACGATGCCAGCACAACTTGAACACTGTTATGCCACTTGTAAAGGAGTTTGAGCA 1230  
QY 1238 gagaataatctccttaccctcgatgactacttgacgggataatgcaatgaatttgcaa 1297  
Db 1231 GCAATGGTTTACCATACATGTTGGTCGACGATATTTCACAGGATTCGGCTTGAATTTGAG 1290  
QY 1298 caattgaaaaatatggc 1314  
Db 1291 CAATTCGGAATATTGC 1307

## RESULT 7

AF139720  
LOCUS Euglena gracilis delta8 fatty acid desaturase (efdl) mRNA, complete cds.  
DEFINITION  
ACCESSION AF139720  
1275 bp mRNA linear PLN 29-JUL-1999



CLQSLFLVLPNGQAHKPSGARVPIISLVQLSLAVHWIYLATMFLFIKDPVNMVMYFL  
VSOAVCGNLLAIVFSLNHNGMPVISKEAVDMDFTKOIIITGRDVHPCGLFANWFTGGL  
NYQIEHHLFPSPRHNFSKIQPAVELCKKGYRVHTTGMIEGTAEVFSRLNEVSKAA  
SKMGKAQ"

BASE COUNT 257 a 395 c 362 g 360 t  
ORIGIN

Query Match 6.7%; Score 98.6; DB 8; Length 1374;  
Best Local Similarity 47.3%; Pred. No. 1.3e-11;  
Matches 416; Conservative 0; Mismatches 439; Indels 24; Gaps 3;

Qy 464 tctgcattgtttattagcacttgatgcacaaattcgatggtttaacacatgattctgc 523  
Db 466 TCGGCTGCCTCTTGGGTCTCTTCGGCAGCAGTGGATGGTGGCGCAGCATTGTTG 525

Qy 524 catcaacagccaaacaaagacaccttgaatgatactattcttcttcttgtaatt 583  
Db 526 CACCACACAGGTCTTCAGGACCGTTTCTGGGGTGATCTTTTCGGCGCCTTCTTGGGAGGT 585

Qy 584 tcttacaaggattttcaagagattggtgaaggaagcaagcataaacactcatcacgtgcc 643  
Db 586 GTCTGCAGGGCTTCGTCTCATGTCATGTCGACATTCACCTCTGTTGACGTGGATGAG 645

Qy 644 acaaatgtattgatcatgacggtgatcatcgacttggcaccacttttcgcattattcca 703  
Db 646 CCCAAGTCCACGGCGAGGATCCCGACATTGACACTCACCTCTGTTGACGTGGATGAG 705

Qy 704 gagatttgcgaagtataagccagcttctgaaaaagcaat-----tctc 748  
Db 706 CATGCTTTGGAGATGTTCTCGGACGTCCTCGACGAGGAGCTGACCGGCATGTGGTGGCGC 765

Qy 749 aagattgtaccatacaacatctctatttcaccgaatgcttccaaatgctcgttttcca 808  
Db 766 TTATGGTCTTAAACAGACCTGGTTCATCTACTTCCCAITCTCTCGTTGCCGCTCTCTCC 825

Qy 809 tggactggctcagtcagttccaatgggtatt-----caaagagaatacaaatggagtaacaag 862  
Db 826 TGGTGCTCCAGTCCATCCTCTTTGTTCTGTCGTAAAGGTCAGGCGCCACAAAGCCCTCTGGA 885

Qy 863 gctcatcaagaataatgcttctggaagcaagcaaatgcttggacacttgggcttgggta 922  
Db 886 GCGCGGTGCCCATTTCCCTTGGTCGAGCAGCTGTCTTTGCCGCTGCACCTGGACCTGGTAC 945

Qy 923 tctatcaattgtcttatttaccacatggccact---tcgggttgcttatttcaattatt 979  
Db 946 CTCGCCACCATGTTCTGTTTCATTAAAGGACCCCGTCAACATGATGGTGACTTTTGGTG 1005

Qy 980 tcacaatggggagggccttttattgctcagtagtcaacttccaacataactctgtt 1039  
Db 1006 TCTCAGGCTGTTTGGCGTAACCTGTGGCGATTGTCTTCGCTCAACCCACACGCTATG 1065

Qy 1040 gataagtatccagcaaatctcgaattttaaacacttgcgcgtcttccaaattttgacc 1099  
Db 1066 CCTGTGATCTCAAGGAGGAAGCCGTCGACATGGATTCTTCACCAAGCAGATCATCAG 1125

Qy 1100 acacgcaacatgactccatctccattcatgattggcttgggtggactcaattatcaag 1159  
Db 1126 GGTGCTGATGTTACCCCTGGTCTGTTCGCCAACTGGTTCACGGGGGATGTGAACCTACCCAG 1185

Qy 1160 atcgagcaccaacttgttcccaacaatgcccagcttgcaatctgaaatgcttgcgtgaaatat 1219  
Db 1186 ATTGACACCACTGTTCTCCCTTCGATGCTCGCCACCAACTTTTCAAAGATCCAGCCTGCT 1245

Qy 1220 gtgaagaatggtgcaagagaataatcttcttcaactcgtcgatgactcaactttgacgga 1279  
Db 1246 GTCGAGACCTTGTGCAAAAAGATGTTGTCGATATCCACACCACTGGCATGATCGAGGGA 1305

Qy 1280 tatgaatgaattgcaacaattgaaaaatatggctgag 1318  
Db 1306 ACTGCAGAGGTCTTTAGCGGTTTGACAGGAGGTCTCCAAG 1344

RESULT 9  
AF306634  
LOCUS  
DEFINITION  
Mortierella isabellina delta6-fatty acid desaturase mRNA, complete cds.  
ACCESSION  
VERSION  
AF306634.1 GI:11559823  
KEYWORDS  
SOURCE  
Mortierella isabellina.  
ORGANISM  
Mortierella isabellina  
Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales;  
Mortierellaceae; Mortierella.  
REFERENCE  
AUTHORS  
1 (bases 1 to 1374)  
Ming-Chun, L., Li, L., Guo-Wu, H., Li, Z. and Lai-Jun, X.  
TITLE  
Cloning and sequencing analysis of delta6-fatty acid desaturase gene from Mortierella isabellina  
JOURNAL  
Junwu Xitong (2001) In press  
REFERENCE  
AUTHORS  
2 (bases 1 to 1374)  
Ming-Chun, L., Li, L., Guo-Wu, H., Li, Z. and Lai-Jun, X.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (18-SEP-2000) Department of Microbiology, Nankai University, 94 Weijin Road, Tianjin 300071, People's Republic of China  
FEATURES  
source  
1..1374  
/organism="Mortierella isabellina"  
/db\_xref="taxon:91625"  
1..1374  
/function="converts linoleic acid to gamma-linolenic acid"  
/note="membrane-bound protein"  
/codon\_start=1  
/product="delta6-fatty acid desaturase"  
/protein\_id="RAG38104.1"  
/db\_xref="GI:11559824"  
/translation="MAAAPSVRTTFRAELINAEALNEGKDAEAFPLMIIDNKVYDVR  
EFVDPHPGGSVILTHVGKDGTDVDFTHPEAAWETLANFYVGDIDSDRAIKNDDFAA  
EVRKLRTLFQSLGYDSSKAYAFKVSFNLCITWGLSTFIVAKWGTSSTLANVLSAALL  
FLWQOCGLAHLDFLHGVFODRPMGLDFLFGVCGVCGSSMMKDKHNTHHAAPNV  
HGEDPDIDTHPLTJWSEHALEMFSDVDPDELTMRMSRFVNLQWTFYFPLLSFARLSW  
CLOSLILLVLPNGQAHKPSGARVISLSVQLSLAHWTWYLATMFLFIKDPVNMVMYFL  
VSOAVCGNLLAIVFSLNHNGMPVISKEAVDMDFTKOIIITGRDVHPCGLFANWFTGGL  
NYQIEHHLFPSPRHNFSKIQPAVELCKKGYRVHTTGMIEGTAEVFSRLNEVSKAA  
SKMGKAQ"

BASE COUNT 260 a 393 c 359 g 362 t  
ORIGIN

Query Match 6.6%; Score 97; DB 8; Length 1374;  
Best Local Similarity 47.2%; Pred. No. 2.9e-11;  
Matches 415; Conservative 0; Mismatches 440; Indels 24; Gaps 3;

Qy 464 tctgcattgtttattagcacttgatgcacaaattcgatggtttaacacatgattctgc 523  
Db 466 TCGGCTGCCTCTTGGGTCTCTTCGGCAGCAGTGGATGGTGGCGCAGCATTGTTG 525

Qy 524 catcaacagccaaacaaagacaccttgaatgatactattcttcttcttgtaatt 583  
Db 526 CACCACACAGGTCTTCAGGACCGTTTCTGGGGTGATCTTTTCGGCGCCTTCTTGGGAGGT 585

Qy 584 tcttacaaggattttcaagagattggtgaaggaagcaagcataaacactcatcacgtgcc 643  
Db 586 GTCTGCCAGGGCTTCGTCTCATGTCATGTCGACATTCACCTCTGTTGACGTGGATGAG 645

Qy 644 acaaatgtattgatcatgacggtgatcatcgacttggcaccacttttcgcattattcca 703  
Db 646 CCCAAGTCCACGGCGAGGATCCCGACATTGACACTCACCTCTGTTGACGTGGATGAG 705

Qy 704 gagatttgcgaagtataagccagcttctgaaaaagcaat-----tctc 748  
Db 706 CATGCTTTGGAGATGTTCTCGGACGTCCTCGACGAGGAGCTGACCGGCATGTGGTGGCGC 765

Qy 749 aagattgtaccatacaacatctctatttcaccgaatgcttccaaatgctcgttttcca 808

```
Db 766 TTCATGGTCTTAAACAGACACCTGGTTCTACTTTCCCATCTCTCTCGTTGGTCCCGCTCTCTCC 825
Qy 809 tggactgggtcagtcagttcacttgaatgggtatt-----caaaagagaatcaaatggagtagacaag 862
Db 826 TGGTGCTCCAGTCCATCTCTCTGTCTGCTTAACGGTCAAGCCCAACAGCCCTCTGGA 885
Qy 863 gtctatcaagaataatgcattctgggagcaagaacaattgttggacattggggttgggta 922
Db 886 GCCCGTGTCTCCATTTCTTGGTTCGAGCAGCTGTCTTGCATGCACCTGGACCTGGTAC 945
Qy 923 tctatcaattgttcttattaccaaatggccact---tcgggtgtctatttctattt 979
Db 946 CTCGCCACCATGTTCTTGTTCAATTAAGGACCCCGTCAACATGATGTGTACTTTTGGTG 1005
Qy 980 tcacaaatgggagggcccttttggatgtcacgtgtcacttgcacttcaaccataactctgtt 1039
Db 1006 TCTCAGGCTGTTTGGCGTAACCTGTTGGCGATTGTGTCTCGCTCAACCCACACAGCGTATG 1065
Qy 1040 gataagatccagcaaatctcgaatttttaacaaacttcgcgcgtcttcaaattttgacc 1099
Db 1066 CCTGTGATCTCCAAGGAGGAAGCCGTCGACATGGATTCTTCAACCAAGCAGATCATCAG 1125
Qy 1100 acacgaacatgactccatctccattcattgattgtcacgtgtcacttgcacttcaaccataactctgtt 1159
Db 1126 GGTGCTGATGTTACACCTCGTCTGTCGCCAATGTGTTACGGGTGGATTTGAATACCAG 1185
Qy 1160 atcgagcacactgtgtcccaacaatggccacgttgcaatctgaatgtcttcgctgaaatat 1219
Db 1186 ATTGAGCACCACTGTTTCCCTTCGATGCTCGCCACAACCTTTTCAAGATCCAGCGTCT 1245
Qy 1220 gtgaaagaatgtgcaagagagaataattcttcaactctgcgtgcgtgacttcttgacgga 1279
Db 1246 GTCGAGACCTTGTGCAAAAGATGTTGTCGATACACCACTGGCATGATCGAGGGA 1305
Qy 1280 tatgcaatgaattgcaacaattgaaaaatatggctgag 1318
Db 1306 ACTGCAGAGGTCTTTAGCCGTTTGAACGAGGTCTCCAAG 1344

RESULT 10
AR080598
LOCUS AR080598 1617 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 1 from patent US 5968809.
ACCESSION AR080598
VERSION AR080598.1 GI:10007328
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Knutzon,D., Mukerji,P., Huang,Y., Thurmond,J. and Chaudhary,S.
TITLE Methods and compositions for synthesis of long chain
poly-unsaturated fatty acids
JOURNAL Patent: US 5968809-A 1 19-OCT-1999;
FEATURES
source Location/Qualifiers
BASE COUNT 310 a 470 c 410 g 427 t
ORIGIN

Query Match 6.6%; Score 97; DB 6; Length 1617;
Best Local Similarity 47.2%; Pred. No. 2.9e-11;
Matches 415; Conservative 0; Mismatches 440; Indels 24; Gaps 3;

Qy 464 tctgcatgtttattagcaacttgatgcaacaattggatgttgaacacatgattctgtc 523
Db 536 TCGGGTGGCTTTGGGCTGTCTTGGCAGCAGTGGGATGGTGGCTCAGCACTTTTGG 595
Qy 524 catcaacagccaacaagaacagaccttgaatgatactattcttcttcttggtaat 583
Db 596 CATCACCAGGTCTTCAGGACCGTTCTGGGGTGATCTTTTCGGCGCCTCTCTTGGGAGGT 655
```

```
Qy 584 ttctacaaggattttcaagagagattggtgaaagcaagcatacaactcatcactgcgtgcc 643
Db 656 GTCTGCCAGGGCTTCTCGTCTGTTGGAAGGCAAGCACAACACTCACCACCGCGCC 715
Qy 644 acaaatgtaattgatcatgacgggtgatatacgaacttggcaccactttttcgattattcca 703
Db 716 CCAACAGTCCACGGCGAGGATCCCGACATTGACACCCACCCTCTGTGTGACCTGGAGTGAG 775
Qy 704 ggagatttggcaagtataaaggccagctttgaaaaagcaat-----tctc 748
Db 776 CATGCCGTTGGAGATGTTCTCGGATGTCCTCCAGATGAGGAGCTGACCCGCGATGGTGGCGT 835
Qy 749 agattgtaccatatatacaacatctctatttcaacgcgaatgcttccaatgetccgttttctca 808
Db 836 TTCATGGTCTGAACACAGACCTGGTTTACTTCCCATCTCTCGTTTGGCCGCTCTCTCC 895
Qy 809 tggactgggtcagtcagttcacttcaatgggtatt-----caaaagagaatcaaatggagtagacaag 862
Db 896 TGGTGCTCCAGTCCATTTCTTGTGCTGCTTAACGGTCAAGCCCAACAGCCCTCGGGC 955
Qy 863 gtctatcaagaataatgcattctgggagcaagaacaattgttggacattggggttgggta 922
Db 956 GCGCGTGTCCCATCTCGTTGGTTCGAGCAGCTGTCTGCGATGCACTGGACCTGGTAC 1015
Qy 923 tctatcaattgttcttattaccaacatggccact---tcgggtgtctatttctattt 979
Db 1016 CTGCCACCATGTTCTCTGTTTCATCAAGGATCCCGTCAACATGCTGTGTACTTTTGGTG 1075
Qy 980 tcacaaatgggagggcccttttggatgtcacgtgtcacttgcacttcaaccataactctgtt 1039
Db 1076 TCGCAGCGGTGTGCGGAAACTTGTGTCGATCGTGTCTGCTCAACCAACAGGTATG 1135
Qy 1040 gataatgtatccagcccaattctgaatttttaacaaacttcgcgcgtcttcaaaatttgacc 1099
Db 1136 CCTGTGATCTCGNAGGAGGCGGTGCGATATGGATTTCTTACGAGCAGATCATCAG 1195
Qy 1100 acacgaacatgactccatctccattcattgattgttcgttgggtggactcaattatcag 1159
Db 1196 GGTGCTGATGTCACCCGGGTCTATTTGGCAACTGTTGTCACGGGTGGATTTGAATCATCAG 1255
Qy 1160 atcgagcacactgtgtcccaacaatggccacgttgcaatctgaatgtcgtgaaatat 1219
Db 1256 ATCGAGCAGCACTGTTTCCCTTCGATGCTCGCCACAACCTTTTCAAGATCCAGCCTGCT 1315
Qy 1220 gtgaaagaatgtgcaagagagaataattcttcaactctgcgtgcgtgacttcttgacgga 1279
Db 1316 GTCGAGACCTTGTGCAAAAGTACATGTCGATACCACTGGCATGATCGAGGGA 1375
Qy 1280 tatgcaatgaattgcaacaattgaaaaatatggctgag 1318
Db 1376 ACTGCAGAGGTCTTTAGCCGTTTGAACGAGGTCTCCAAG 1414

RESULT 11
AR098439
LOCUS AR098439 1617 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6075183.
ACCESSION AR098439
VERSION AR098439.1 GI:12807696
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Knutzon,D., Mukerji,P., Huang,Y., Thurmond,J. and Chaudhary,S.
TITLE Methods and compositions for synthesis of long chain
poly-unsaturated fatty acids in plants
JOURNAL Patent: US 6075183-A 1 13-JUN-2000;
FEATURES
source Location/Qualifiers
BASE COUNT 310 a 470 c 410 g 427 t
ORIGIN
```

Query Match									
Best Local Similarity 6.6%; Score 97; DB 6; Length 1617;									
Matches 415; Conservative 0; Mismatches 440; Indels 24; Gaps 3;									
QY	464	tctgcatgtttattagcaacttgcacgaattggatggttaacacatgagttctgc	523						
DB	536	TCGGCTGCGCTTTTGGGTCTGTTCTGGCAGCAGTGGGATGTTGGCTCAGCACTTTTGG	595						
QY	524	catcaacagccaacaaagacagacctttgaatgatactatttctttgtttgtgta	583						
DB	596	CATCACAGGCTTCCAGGACCGTTCTGGGGTGATCTTTTCGGGCGCTTCTTGGGAGGT	655						
QY	584	ttctacaaggattttcaagagattggtgaaggagcaagataaacactcatcagctgc	643						
DB	656	GTCCTGCCAGGGCTTCTGCTCCCTGGTGAAGGACAGCACAACACTCACCGCGGC	715						
QY	644	acaaatgtaattgatcatgacggtgatatcgcacttggcaccacttttgcattttcca	703						
DB	716	CCCAAGTCCACGGCGAGGATCCCGACATTTGACACCCACCTCTGTTGACCTGGAGTGAG	775						
QY	704	gagatttgtcaagataaagccagctttgaaaaagcaat-----tctc	748						
DB	776	CATGCGTTGGAGATGTTCTCGGATGTCCAGATGAGGAGCTGACCGCGCATGTGCGCGT	835						
QY	749	aagattgtaccatacaacatctctatttcacgcaaatgcttccaatgctccgtttctca	808						
DB	836	TTCAATGGTCTGAACGAGACCTGGTTTACTTCCCATTTCTCTGTTGCCGCTCTCTCC	895						
QY	809	tggactggctcagtcagttcaatgggtatt-----caaaagaaatacaaatggagta	862						
DB	896	TGGTGCTCCAGTCCATTTCTTTGTGCTGCTTACCGGTCAAGCCCTCGGCG	955						
QY	863	gtctatcaaaagaatgcttgggagcaagcaaatgttggacattgggcttgggta	922						
DB	956	GCGCGTGTGCCATCTCGTTGTCGAGCAGCTGCTGTCGATGAGGAGCTGACCGCATGTGCGCGT	1015						
QY	923	ttctatcaaatgttcttattaccacatggccact---tcgggttgcattattcatt	979						
DB	1016	CTCGCCACCATGTTCTCTGTTTCATCAAGGATCCCGTCAACATGCTGTGTTACTTTTGGTG	1075						
QY	980	tcacaaatggaggagccttttgcacagtaglcaacttcaaccataactctgt	1039						
DB	1076	TGCGAGCGGTGCGGAAACTTTGGCGATCTGTTCTCGCTCAACCAACGATG	1135						
QY	1040	gataagatccagcaaatctctogaattttaaacacttgcgcgtcttcaaat	1099						
DB	1136	CCTGTGATCTCAAGGAGGCGGTTCGATATGGATTTCTTCCAGAACGAGATCATCAG	1195						
QY	1100	acacgcaacatgactccatctccattcattgattggcttgggtggactcaattcag	1159						
DB	1196	GGTCGTGATGTCCACCGGGTCTATTTCGCAACTGTTTACGGGTGGATTGAACATATCAG	1255						
QY	1160	atcgagcaccacttgttcccaacaatggcagcttgaactgttcgttgaata	1219						
DB	1256	ATCGAGCACCATTGTTCCCTTCGATGCTCCGCCACAATTTTCAAAGATCCAGCTGCT	1315						
QY	1220	gtgaagaatggtgcaagagaaatacttcttaccctgtcgatgactactttgacgga	1279						
DB	1316	GTCGAGACCTGTGCAAAAGTACATGTCCGATACCAACACCGGTATGATCGAGGGA	1375						
QY	1280	tatgcaatgaaattgcaacaattgaaaaatagcgtag	1318						
DB	1376	ACTGCAGAGGCTTTAGCGCTGTAACGAGGTCTCAAG	1414						
RESULT 12									
AR136018									
LOCUS AR136018 1617 bp DNA linear PAT 16-JUN-2001									
DEFINITION Sequence 1 from patent US 6136574.									
ACCESSION AR136018									
VERSION AR136018.1 GI:14476690									

KEYWORDS									
SOURCE Unknown.									
ORGANISM Unknown.									
REFERENCE 1 (bases 1 to 1617)									
AUTHORS Knutson,D., Mukerji,P., Huang,Y., Thurmond,J. and Chaudhary,S.									
TITLE Methods and compositions for synthesis of long chain polyunsaturated fatty acids									
JOURNAL Patent: US 6136574-A 1 24-OCT-2000;									
FEATURES Location/Qualifiers									
source 1..1617									
BASE COUNT 310 a 470 c 410 g 427 t									
ORIGIN /organism="unknown"									
Query Match 6.6%; Score 97; DB 6; Length 1617;									
Best Local Similarity 47.2%; Pred. No. 2.9e-11;									
Matches 415; Conservative 0; Mismatches 440; Indels 24; Gaps 3;									
QY	464	tctgcatgtttattagcaacttgcacgaattcgatggttaacacatgagttctgc	523						
DB	536	TCGGCTGCGCTTTTGGGTCTGTTCTGGCAGCAGTGGGATGTTGGCTCAGCACTTTTGG	595						
QY	524	catcaacagccaacaaagacagacctttgaatgatactatttctttgtttgtgta	583						
DB	596	CATCACAGGCTTCCAGGACCGTTTCGGGGTGATCTTTTCGGGCGCTTCTTGGGAGGT	655						
QY	584	ttctacaaggattttcaagagattggtgaaggagcaagataaacactcatcagctgc	643						
DB	656	GTCCTGCCAGGGCTTCTGCTCCCTGGTGAAGGACAGCACAACACTCACCGCGGC	715						
QY	644	acaaatgtaattgatcatgacggtgatatcgcacttggcaccacttttgcattttcca	703						
DB	716	CCCAAGTCCACGGCGAGGATCCCGACATTTGACACCCACCTCTGTTGACCTGGAGTGAG	775						
QY	704	gagatttgtcaagataaagccagctttgaaaaagcaat-----tctc	748						
DB	776	CATGCGTTGGAGATGTTCTCGGATGTCCAGATGAGGAGCTGACCGCGCATGTGCGCGT	835						
QY	749	aagattgtaccatacaacatctctatttcacgcaaatgcttccaatgctccgtttctca	808						
DB	836	TTCAATGGTCTGAACGAGACCTGGTTTACTTCCCATTTCTCTGTTGCCGCTCTCTCC	895						
QY	809	tggactggctcagtcagttcaatgggtatt-----caaaagaaatacaaatggagta	862						
DB	896	TGGTGCTCCAGTCCATTTCTTTGTGCTGCTTACCGGTCAAGCCCTCGGCG	955						
QY	863	gtctatcaaaagaatgcttgggagcaagcaaatgttggacattgggcttgggta	922						
DB	956	GCGCGTGTGCCATCTCGTTGTCGAGCAGCTGCTGTCGATGAGGAGCTGACCGCATGTGCGCGT	1015						
QY	923	ttctatcaaatgttcttattaccacatggccact---tcgggttgcattattcatt	979						
DB	1016	CTCGCCACCATGTTCTGTTTCATCAAGGATCCCGTCAACATGCTGTGTTACTTTTGGTG	1075						
QY	980	tcacaaatggaggagccttttgcacagtaglcaacttcaaccataactctgt	1039						
DB	1076	TGCGAGCGGTGCGGAAACTTTGGCGATCGTGTCTCGCTCAACCAACGATG	1135						
QY	1040	gataagatccagcaaatctctogaattttaaacacttgcgcgtcttcaaat	1099						
DB	1136	CCTGTGATCTCAAGGAGGCGGTTCGATATGGATTTCTTCCAGAACGAGATCATCAG	1195						
QY	1100	acacgcaacatgactccatctccattcattgattggcttgggtggactcaattcag	1159						
DB	1196	GGTCGTGATGTCCACCGGGTCTATTTCGCAACTGTTTACGGGTGGATTGAACATATCAG	1255						
QY	1160	atcgagcaccacttgttcccaacaatggcagcttgaactgttcgttgaata	1219						
DB	1256	ATCGAGCACCATTGTTCCCTTCGATGCTCCGCCACAATTTTCAAAGATCCAGCTGCT	1315						
QY	1220	gtgaagaatggtgcaagagaaatacttcttaccctgtcgatgactactttgacgga	1279						
DB	1316	GTCGAGACCTGTGCAAAAGTACATGTCCGATACCAACACCGGTATGATCGAGGGA	1375						
QY	1280	tatgcaatgaaattgcaacaattgaaaaatagcgtag	1318						
DB	1376	ACTGCAGAGGCTTTAGCGCTGTAACGAGGTCTCAAG	1414						

Db	1316	GTGAGACCTGTGCAAAAGTACAAATGCCGATACACACACCGGTATGATCGAGGGA	1375
QY	1280	tatgcaatgaattgcaacaattgaaaaatatggctgag	1318
Db	1376	ACTGCAGAGTCTTAGCGCTGTAACGAGGCTCTCAAG	1414
RESULT 13			
LOCUS	AF110510	1617 bp	linear
DEFINITION	Mortierella alpina delta-6 fatty acid desaturase mRNA, complete cds.	PLN 18-NOV-1999	
ACCESSION	AF110510		
VERSION	AF110510.1	GI:6448795	
KEYWORDS			
SOURCE	Mortierella alpina.		
ORGANISM	Mortierella alpina		
REFERENCE			
AUTHORS	Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mortierellaceae; Mortierella.		
TITLE	Huang, Y.S., Chaudhary, S., Thurmond, J.M., Bobik, E.G. Jr., Yuan, L., Chan, G.M., Kirchner, S.J., Mukerji, P. and Knutzon, D.S. Cloning of delta12- and delta6-desaturases from Mortierella alpina and recombinant production of gamma-linolenic acid in Saccharomyces cerevisiae		
JOURNAL	Lipids 34 (7), 649-659 (1999)		
MEDLINE	99406036		
PUBMED	10478922		
REFERENCE	2 (bases 1 to 1617)		
AUTHORS	Knutzon, D.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-DEC-1998) Calgene LLC, 1920 Fifth St., Davis, CA 95616, USA		
FEATURES			
source	1. 1617		
	/organism="Mortierella alpina"		
	/strain="ATCC32221"		
	/db_xref="ATCC:32221"		
	/db_xref="taxon:64518"		
	/clone="Ma524"		
	71..1444		
CDS	/codon_start=1		
	/product="delta-6 fatty acid desaturase"		
	/protein_id="AAF08685.1"		
	/db_xref="GI:6448796"		
	/translation="MAAAPSVRTFTFAEVLNAELNEGKDAEAPFLMIDNKVYDVR EFVDPHGSVILTHVGKGDVDFTHPEAAWETLANEYVGDIDESDRDIKNDDFAA EYRKLRTFLQSLGYDSSKAYAFKVSFNLCIWLGLSTVIVAKGOTSLANVLSAALL GLFWQCCGLAHDFLIHQVDFQDFWGLDFGLGVCVCGQSSWWKDKHNTTHAAPNV HGEPDIDTHPLLWSEHALEMFSDVPEELTRMSRFVNLQWTFYFPIISFARLSW CLOSLFLVLPNGQAKPSGARVPISLVQLSLAMHWTLATMELFKIDPVNMLVYFL VSOAVCGNLLAIVFSLNNGMPVISKEEAVDMDFTKOITITGRDVHPGLFANWFTGGL NYQIEHHFLFSPMRHNFPSKIQPAVELLCKKNVRYHTTGMIEGTAEVSRLENSKAA SKMGKAQ"		
BASE COUNT	310 a 470 c 410 g 427 t		
ORIGIN			
Query Match	6.6%; Score 97; DB 8; Length 1617;		
Best Local Similarity	47.2%; Pred. No. 2.9e-11;		
Matches 415; Conservative	0; Mismatches 440; Indels 24; Gaps 3;		
QY	464	tctgcatgtttattagcacttgatgcaacattcgatgggttaacacatgattctctgc	523
Db	536	TCGGTGCCTTTTGGTCTGTCTGGCAGCAGTGGGATGCTGGCTCAGCACTTTTG	595
QY	524	catcaacagccacaagaacagaccttgaatgatatactattctttgttttggaat	583
Db	596	CATCACCAGGTCTTCAGACCGCTTCTGGGTGATCTTTTCGGCGCCTCTCTGGGAGGT	655
QY	584	tctctacaaggattttcaagagattggtggaagacaagacaataactcatacagctgcc	643

Db	656	GTCTGCAGGGCTTCGTCTCTCGTGGTGAAGGACAAGCACAACTCACCACGCGGCC	715
QY	644	acaaatgaattgatcatgacgggtgatcgcacttggcaccacttttgcattattcca	703
Db	716	CCCAACGTCCACGGCGAGGATCCGACATGACACCCACCCTCTGTTGACCTGGAGTGAG	775
QY	704	ggagatttgcagataaaggccagcttttgaanaagcaat-----tctc	748
Db	776	CATGCGTTGGAGATGTTCTCGGATGTCCCGATGAGGAGCTGACCGCATGTGGTGGCGT	835
QY	749	agagattgaccatatcaaatctctatttccacgcgaatgcttccaatgctccgctttctca	808
Db	836	TTCAATGCTGCTGAACAGACCTGCTTTTACTTCCCATTTCTCTGTTTGGCCCTCTCTCC	895
QY	809	tggactggctcagtcagttcaatgggtatt-----caaaagagaatcaaatggagtagacaag	862
Db	896	TGGTGCCTCCAGTCCATTCTCTTTTGTGCTTAACTGCTCAGGCGCCACAAAGCCTCGGGC	955
QY	863	gtctatcaaaagaatgcattcttgggaggaagaacaattgttggacattggggttgggta	922
Db	956	GCAGCGTGTGCCATCTCTGTTGGTGCAGACGCTGCTGCTTGGATGACACTGGACCTGGTAC	1015
QY	923	tctatcaatigtcttatttaccacaatggcact---tcgggttcttatttatttatt	979
Db	1016	CTCGCCACCATGTTCTCTGTCATCAAGGATCCCGTCAACATGCTGGTGTACTTTTGGTG	1075
QY	980	tcacaaatgggagggccttttattgtctcagctagtcactttcaaccaataactctgtt	1039
Db	1076	TCGACGCGGTGTGCGGAACCTTTGTTGGCGATGCTGTCTCTCGCTCAACCAACAGGTATG	1135
QY	1040	gataaglatccagccaattctctgaattttaacaacttcgcctctcttcaaatatttgacc	1099
Db	1136	CCTGTGATCTCGAAGGAGGCGGCTCGATATGATGATTTCTTCACGAAGCAGATCATCAG	1195
QY	1100	acacgaacatgactcctctccattcattgattgctcagctagtcactttcaaccaataactca	1159
Db	1196	GGTGTGATGTCACCGCGGTCTATTTGGCACTGTTCACGGGTGATTTGAATACTATCAG	1255
QY	1160	atcgagcaccacttgttcccaacaatgccacgttgcaactctgaatgcttgcgtgaaatat	1219
Db	1256	ATCGAGCACCATTTGTTCCCTTCGATGCTCGCCCAACTTTTCAAAGATCCAGCCTGCT	1315
QY	1220	gtgaaagaattgtgcaaaagagaataatcttcttactactcgtcgatgactactttgacga	1279
Db	1316	GTGAGACCGCTGTGCAAAAGTACAATGTCCGATACACACACCGGTATGATCGAGGA	1375
QY	1280	tatgcaatgaattgcaacaattgaaaaatatgctgag	1318
Db	1376	ACTGCAGAGTCTTAGCGCTCTGAACGGGTCTCCAAG	1414
RESULT 14			
LOCUS	CET13F2	38060 bp	DNA linear
DEFINITION	Caenorhabditis elegans cosmid T13F2, complete sequence.	INV 11-DEC-2001	
ACCESSION	Z81122		
VERSION	Z81122.1	GI:1628203	
KEYWORDS	HTG; Caveolin; Cyclin D; Cytochrome B5; Human KIAA0170 protein; Major Sperm Protein; Single-stranded DNA binding protein p9. Caenorhabditis elegans.		
SOURCE	Caenorhabditis elegans.		
ORGANISM	Caenorhabditis elegans		
REFERENCE	1 (sites)		
AUTHORS	none.		
TITLE	Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium Science 282 (5396), 2012-2018 (1998)		
JOURNAL	99069613		
MEDLINE	The C. elegans Sequencing Consortium.		
REMARK	2 (bases 1 to 38060)		
REFERENCE	Swinburne, J.		
AUTHORS	Direct Submission		
TITLE			



## JOURNAL

Submitted (21-OCT-1996) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jes@sanger.ac.uk or rwnematode.wustl.edu

## COMMENT

Coding sequences below are predicted from computer analysis, using predictions from GeneFinder (P. Green, U. Washington), and other available information.

Current sequence finishing criteria for the *C. elegans* genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones.

Exceptions are indicated by an explicit note.

For a graphical representation of this sequence and its analysis see: <http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=Tl3F2>

**IMPORTANT:** This sequence is not the entire insert of clone Tl3F2. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

The true left end of clone Tl3F2 is at 1 in this sequence. The true left end of clone W08D2 is at 37953 in this sequence. The start of this sequence (1..104) overlaps with the end of sequence AL022473. The end of this sequence (37953..38060) overlaps with the start of sequence 270271.

**IMPORTANT:** This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

## FEATURES

## source

Location/Qualifiers

1..38060

/organism="Caenorhabditis elegans"

/db\_xref="taxon:6239"

/chromosome="IV"

/clone="Tl3F2"

complement(join(616..840,1149..1247))

/gene="Tl3F2.12"

complement(join(616..840,1149..1247))

/note="contains similarity to pfam domain: PF00635 (MSP

(Major sperm protein) domain), Score=115.5,

E-value=3.3e-31, N=1

cdna EST yk305b4.3 comes from this gene"

/codon\_start=1

/protein\_id="CAB03363.1"

/db\_xref="GI:3879839"

/db\_xref="SPTREMBL:Q94054"

/translation="MLRIEPPSATFPASGSGSTHTITSVNESRMAFKVKSSNNEHYRV  
RPVIGFVEARGMKFEIRLEGPVKDDKIMLOYAEVPADETDAQAPKAGAQGDVTI  
LLKTT"

complement(1681..2064)

/gene="Tl3F2.10"

complement(1681..2064)

/gene="Tl3F2.10"

/note="contains similarity to pfam domain: PF00635 (MSP

(Major sperm protein) domain), Score=117.8,

E-value=6.8e-32, N=1

cdna EST yk55e9.3 comes from this gene

cdna EST yk55c3.5 comes from this gene

cdna EST yk204d3.5 comes from this gene

cdna EST yk235f10.5 comes from this gene

cdna EST yk261h12.3 comes from this gene

cdna EST yk261h12.5 comes from this gene

cdna EST yk263e5.5 comes from this gene

cdna EST yk381d1.5 comes from this gene"

/codon\_start=1

/protein\_id="CAB03361.1"

/db\_xref="GI:3879837"

/db\_xref="SPTREMBL:Q9TVW5"

/translation="MAQSVPPGDIQTQGTQKIVFNAPYDDKHTYHIKVINSSARRIGY  
GIKTTNMRKLGVDPPCGVLDPKAEVLLAVSCDAFAFGQEDTNNDRITVENTPTDGA  
KQRRWFQDGMVRRKNLPIEYNP"

join(2967..3107,3153..3272,3388..3894)

/gene="Tl3F2.9"

## CDS

join(2967..3107,3153..3272,3388..3894)  
/gene="Tl3F2.9"  
/note="cdna EST EMBL:AW057187 comes from this gene"  
/codon\_start=1  
/protein\_id="CAB03360.1"  
/db\_xref="GI:3879836"  
/db\_xref="SPTREMBL:Q94052"  
/translation="MVKKITVTAQGFLEMIERQAEORRETVPVLCPIVEKAQPRTA  
LNKQSCPVPTTARVTEIEKKSISCPALLHADOTIEMALSPKVNGLFAASSTSTA  
SLFKYGGFEKTEVATEDVNTFFWNSPAHSPQLPVFKYSPIDSKKIYCAVHNFN  
ISEDKYFFESSFFKYLILLETPTAOSVSSPISPSSTSSSETIILPGHLLQE  
VYLQRAVIEEQKKIIDSFHEIQKLNKLAQNFTA"  
4956..5339  
/gene="Tl3F2.11"  
4956..5339  
/gene="Tl3F2.11"  
/note="contains similarity to pfam domain: PF00635 (MSP

## gene

## CDS

(Major sperm protein) domain), Score=116.7,  
E-value=1.4e-31, N=1  
cdna EST yk94h6.3 comes from this gene  
cdna EST yk124b1.3 comes from this gene  
cdna EST yk99g7.5 comes from this gene  
cdna EST yk202d5.3 comes from this gene  
cdna EST yk202d5.5 comes from this gene  
cdna EST yk223g7.3 comes from this gene  
cdna EST yk223g7.5 comes from this gene"  
/codon\_start=1  
/protein\_id="CAB03362.1"  
/db\_xref="GI:3879838"  
/db\_xref="SPTREMBL:Q94053"  
/translation="MAQSVPPGDIQTQGTQKIVFNAPYDDKHTYHIKVINSSARRIGY  
GIKTTNMRKLGVDPPCGVLDPKAEVLLAVSCDAFAFGQEDTNNDRITVENTPTDGA  
KQRRWFQDGMVRRKNLPIEYNP"  
join(7442..7717,7843..8274)  
/gene="Tl3F2.8"  
join(7442..7717,7843..8274)  
/gene="Tl3F2.8"  
/note="contains similarity to pfam domain: PF01146

## gene

## CDS

(Caveolin), Score=378.4, E-value=2.3e-110, N=1  
cdna EST yk51c7.3 comes from this gene; cdna EST yk61c3.3  
comes from this gene  
cdna EST yk61d4.3 comes from this gene; cdna EST yk61h12.3  
comes from this gene  
cdna EST yk62f1.3 comes from this gene; cdna EST yk64e1.3  
comes from this gene  
cdna EST yk74b2.3 comes from this gene; cdna EST yk78g5.3  
comes from this gene  
cdna EST yk78h8.3 comes from this gene; cdna EST yk80h11.3  
comes from this gene  
cdna EST yk96a9.3 comes from this gene; cdna EST yk96c2.3  
comes from this gene  
cdna EST yk98b3.3 comes from this gene; cdna EST yk103d6.3  
comes from this gene  
cdna EST yk105g5.3 comes from this gene; cdna EST  
yk105h10.3 comes from this gene  
cdna EST yk106d8.3 comes from this gene; cdna EST  
yk106f7.3 comes from this gene  
cdna EST yk106h8.3 comes from this gene; cdna EST  
yk112g9.3 comes from this gene  
cdna EST yk51c7.5 comes from this gene; cdna EST yk61c3.5  
comes from this gene  
cdna EST yk61d4.5 comes from this gene; cdna EST yk61h12.5  
comes from this gene  
cdna EST yk62f1.5 comes from this gene; cdna EST yk64e1.5  
comes from this gene  
cdna EST yk74b2.5 comes from this gene; cdna EST yk78g5.5  
comes from this gene  
cdna EST yk78h8.5 comes from this gene; cdna EST yk80h11.5  
comes from this gene  
cdna EST yk96a9.5 comes from this gene; cdna EST yk96c2.5  
comes from this gene  
cdna EST yk98b3.5 comes from this gene; cdna EST yk103d6.5  
comes from this gene



cdna EST yk105q5.5 comes from this gene; cdna EST yk105h10.5 comes from this gene  
cdna EST yk106d8.5 comes from this gene; cdna EST yk106f7.5 comes from this gene  
cdna EST yk106h8.5 comes from this gene; cdna EST yk112g9.5 comes from this gene  
cdna EST yk108d10.3 comes from this gene; cdna EST yk108d10.5 comes from this gene  
cdna EST yk82b11.5 comes from this gene; cdna EST yk90f7.5 comes from this gene  
cdna EST EMBL:D75025 comes from this gene; cdna EST yk468f2.3 comes from this gene  
cdna EST yk468f2.5 comes from this gene; cdna EST yk293b3.3 comes from this gene  
cdna EST yk293b3.5 comes from this gene; cdna EST yk289b3.3 comes from this gene  
cdna EST yk289b3.5 comes from this gene; cdna EST yk291h7.3 comes from this gene  
cdna EST yk291h7.5 comes from this gene; cdna EST yk241g9.3 comes from this gene  
cdna EST yk241g9.5 comes from this gene; cdna EST yk597g3.3 comes from this gene  
cdna EST yk610g6.3 comes from this gene; cdna EST yk612a3.3 comes from this gene  
cdna EST yk612a6.3 comes from this gene; cdna EST yk612e6.3 comes from this gene  
cdna EST yk665d1.3 comes from this gene; cdna EST yk665f11.3 comes from this gene  
cdna EST yk666h10.3 comes from this gene; cdna EST yk597g3.5 comes from this gene  
cdna EST yk610g6.5 comes from this gene; cdna EST yk612a3.5 comes from this gene  
cdna EST yk612a6.5 comes from this gene; cdna EST yk612e6.5 comes from this gene

Query Match 6.6%; Score 96.2; DB 3; Length 38060;  
Best Local Similarity 55.6%; Pred. No. 3.6e-11;  
Matches 185; Conservative 0; Mismatches 148; Indels 0; Gaps 0;  
QY 263 gttgatcaatgtatcagcatatgatgtcagtggtgcagagaaagaaatggtgaa 322  
DB 35193 GATGATGTCGAATGGCACTTTCATATATTCGACAAAGGATCTCCCAATAATATAA 35252  
QY 323 tcaatgaaactacgacagagcttcgatgatgatgattgaatgaagcaaatgaaaca 382  
DB 35253 AGTTTCACTGATACGTATGCGAGTTCTGCGAGAAGGACTTATGGATGATCTCTTTG 35312  
QY 383 tatttctgttaagcggattcaaacacttcaattatgacattgacattgatttattctcag 442  
DB 35313 TTCTACATTTAGAAAAATCTTGAACAAATCTTCACAAATCTTTTTCATTTACCTTCAA 35372  
QY 443 tatcttgatggtatattacttctgcattgttattagcacttgcatgagcaacaattcgaa 502  
DB 35373 TACCACACATATTAATCTCCATCAGCTATCTAATGGGAGTTGCGTGGCAAAATTTGGGA 35432  
QY 503 tggttaacacatgagttctgcatcatcacgcaacaaagacacgaccttggatgatact 562  
DB 35433 TGGTTAATCCATGAATCCGACATCATCATAGTTGTTCAAAAACAGATACATATGATTG 35492  
QY 563 atttcttcttcttggttaatttcttacaagga 595  
DB 35493 GCCAGCTATTTCGTGGAAACTTTTTCACAAGTA 35525

RESULT 15  
AF307940  
LOCUS  
DEFINITION  
Mortierella alpina clone pTMACL-3 delta 6-fatty acid desaturase  
mRNA, complete cds.  
ACCESSION  
AF307940  
VERSION  
AF307940.1 GI:12007241  
KEYWORDS  
Mortierella alpina.

ORGANISM Mortierella alpina  
Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales;  
Mortierellaceae; Mortierella.  
REFERENCE 1 (bases 1 to 1374)  
AUTHORS Xing,L., Li,M., Liu,L., Hu,G. and Zhang,L.  
TITLE Expression of Mortierella alpina delta 6-fatty acid desaturase gene in Saccharomyce cerevisiae  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1374)  
AUTHORS Xing,L., Li,M., Liu,L., Hu,G. and Zhang,L.  
TITLE Direct Submission  
JOURNAL Submitted (26-SEP-2000) Department of Microbiology, Nankai University, 94 Weijin Road, Tianjin 300071, P.R. China  
FEATURES  
Location/Qualifiers  
1..1374  
/organism="Mortierella alpina"  
/db\_xref="taxon:64518"  
/clone="pTMACL-3"  
1..1374  
/codon\_start=1  
/product="delta 6-fatty acid desaturase"  
/protein\_id="AAG45092.1"  
/db\_xref="GI:12007242"  
/translation="MAAAPSVRTFTTAEILNBAALNEGKKDAEAPFLMIDNKKYDVR  
EFVDPHPGGSVILTHVKGDTDFVDFTHPEAAWETLANFYVGDIDESRAIKNDPFAA  
EVRKLRTLFQSLGYDYDSSKAYYAFKVFNLICWGLSTFIVAKRGQSTLANELSAALL  
GLFWOORGLAHDFLHHOYFQDRFGWDLFGALFGDCQCFGSSWWKDKHNTHAAPNV  
HGCDPDIDTHPLLTTHSEHALEMESDVDPDELTMRMSRFMLNQWTFYFPLSFARLSW  
CLOSILFVLPNGQAHKPSGARVPISLVEOLSLAMHTWTLATMFLFIKDPVNMVYFL  
VSOAVCGNLLAIVFSLNHNMPVISKEAVDMDFFTKQIITGRDVRHPLGFANWFTGGGL  
NYQIEHHLFPSPMRHNFSKIQPAVETLCKKYGVRYHTTGMIEGTAEVFSRLNEVSKAA  
SRMGRAQ"  
BASE COUNT 266 a 395 c 361 g 352 t  
ORIGIN  
Query Match 6.2%; Score 90.6; DB 8; Length 1374;  
Best Local Similarity 46.8%; Pred. No. 7.6e-10;  
Matches 411; Conservative 0; Mismatches 444; Indels 24; Gaps 3;  
QY 464 tctgcatgtttattagcaattgcattggcaacaattcggatgggttaacacatgattctgc 523  
DB 466 TCGGCTCGCGCTCTTGGGTCTCTTCTGGCAGCAGAGAGATGGTTGGCGCAGACTTTTG 525  
QY 524 catcaacagcacaacaagaacagacctttgaatgatactattcttcttcttggtaat 583  
DB 526 CACCACCAGGCTCTTCAGGACCGCTTTCGGGAGAGACTTTTCGGCGCCCTCTCTGGGAGGA 585  
QY 584 ttcttcaaggatttttcaagagattggtggaagcaagcataacacatcatcagctgcc 643  
DB 586 GACTGCGAGGGCTCTCGTCTCATGTTGGAAGGACAAGCACAACACTCACCACGCGCC 645  
QY 644 acaaatgtaattgatcatgacggtgatatactgacttggcaccacttttcgacttattcca 703  
DB 646 CCCAACGTCCACGGCAGGATCCCGACATTGACACTCACCTCTGTGTGAGTGGAGTGAG 705  
QY 704 ggagattgtgcaagtataaagccagctttgaaaaagcaat-----tctc 748  
DB 706 CATGCTTTGGAGATGTTCTCGGACGTCCTTGACGAGAGCTGACCCGCGATGTGGTCGCGC 765  
QY 749 aagattgtaccatatcaacatctctatttcacgcgaatgcttccaatgctccgtttctca 808  
DB 766 TTCATGGTCTTAACCCAGACCTGGTTCCTACTTCCCATTTCTCTCGTTGGCCGCTCTCC 825  
QY 809 tggactgggtcagtcagtccaatgggtatt-----caaaagagaatcaaatggagttacaag 862  
DB 826 TGGTGGCTCCAGTCCCATCTCTTTTGTCTGCCCTAACGGTCAGGCCCAACAGCCCTCTGGA 885  
QY 863 gtcttcaaaagaatacattcttgggagcaagcaacaatttcttgacattggcgttggtga 922  
DB 886 GCCCGTGTGCCCATTTCCCTTGGTCGAGCAGCTGCTCTCTGCCATGACACTGGTAC 945  
QY 923 ttctatcaattgttcttattaccacaacatggccact---tcggggtgtcttatttctatt 979

```

      || ||||| || | || | | | | | | | |
Db 946 CTCGCCACCATGTTCTTGTTCATTAAGGACCCCGTCAACATGATGGTGACTTTTGGTG 1005
QY 980 tcacaaatggaggagcccttttgattgctcacgtagtcactttcaaccataactctgtt 1039
      || || | | | | | | | | | | | | | |
Db 1006 TCTCAGGCTGTTTGGCGTAAACCTGTGGCGATGTGTTCTCGCTCAACCACACACGGGTATG 1065
QY 1040 gataagtatccagcccaattctcgaattttaacaacacttcgcgcgtctcttcaaaattttgacc 1099
      | | | | | | | | | | | | | | | | | |
Db 1066 CCTGTGATCTCCAAGGAGGAAGCGTCGACATGGATTCTTCACCAAGCAGATCATCAG 1125
QY 1100 acacgcaaatgactccatccattcattgattggtcttggggtggaactcaattatcag 1159
      || | | | | | | | | | | | | | | | |
Db 1126 GFTCGTGATGTTTCACCTGGTCTGTTTCGCCAACTGGTTCACGGGTGGATTGAACCTACCAG 1185
QY 1160 atcgagcacactgttcccaacaatgccacgttgcaatcgaatgcttgctgaaatat 1219
      || ||||| ||||| ||||| || ||||| || |||||
Db 1186 ATTGACACCACTTGTTCCTTCGATGCCCTCGCCACAACCTTTTCAAAAGATCCAGCCTGCT 1245
QY 1220 gtgaaagaatggtgcaaaagagaataatcttacctcgtcgtgatgactactttgacgga 1279
      || | | | | | | | | | | | | | | | |
Db 1246 GTCGAGACCTTGTGCAAAAAGTATGTTCCGATACCAACACCACTGGCATGATCGAGGGA 1305
QY 1280 tatgcaatgaattgcaacaattgaaaaatatggctgag 1318
      |||| | | | | | | | | | | | |
Db 1306 ACTGCAGAGGCTCTTATGCCGCTTTGAACGAGGTCTCCAAG 1344
```

Search completed: August 7, 2002, 02:55:26  
Job time: 5057 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 01:39:59 ; Search time 272.14 Seconds  
(without alignments)  
9223.668 Million cell updates/sec

Title: US-09-555-093-1  
Perfect score: 1462  
Sequence: 1 gctcaccaaatgtgtcg.....ataaaacctgttttcaa 1462

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1451	99.2	1463	20	AAAX76589
2	350.4	24.0	1461	21	AAAS1232
3	348.2	23.8	1344	20	AAAX86961
4	112.8	7.7	1275	21	AAAS1233
5	97	6.6	1374	21	AAZ47129
6	97	6.6	1374	22	AAZ47129
7	97	6.6	1617	19	AAAF63624
8	97	6.6	1617	20	AAAX00889
9	97	6.6	1617	21	AAA09430

10	76.2	5.2	2012	22	AAF26040	P. patens delta6-d
11	68.4	4.7	1684	19	AAV34398	Borage delta-6 des
12	68.4	4.7	1684	20	AAX24917	Borage delta-6 des
13	68.4	4.7	1685	17	AAT30395	Borage delta-6-des
14	62	4.2	1467	22	AAF25730	C. purpureus delta
15	62	4.2	2040	22	AAF23729	C. purpureus delta
16	58.2	4.0	1471	21	AAAD01349	Florida bitterbush
17	56.2	3.8	1650	21	AAC33846	Arabidopsis thalia
18	56.2	3.8	1650	21	AAC51462	Arabidopsis thalia
19	56.2	3.8	1678	21	AZ44833	A. thaliana sldl D
20	55.8	3.8	2160	22	AAF25734	C. purpureus delta
21	53.6	3.7	1972	21	AAAD01353	Wheat sphingolipid
22	52	3.6	494	19	AAV63634	Phaeodactylum tric
23	52	3.6	494	20	AAV63634	M. alpina desatura
24	52	3.6	494	20	AAV82633	DNA encoding a put
25	52	3.6	494	21	AAV82633	P. tricornutum fat
26	52	3.6	494	21	AAA14585	DNA encoding a Pha
27	50.2	3.4	1606	21	AAZ44851	Sphingolipid desat
28	47.8	3.3	1934	21	AAAD01352	Soybean sphingolip
29	47.4	3.2	1594	21	AAZ44832	B. napus sldl DNA.
30	47	3.2	1335	22	AAI66599	Rat fatty acid des
31	47	3.2	1335	22	AAI66599	Rat delta-6-desatu
32	46.2	3.2	880	21	AAAD01351	Soybean sphingolip
33	46	3.1	2943	21	AAV70229	Plasmodium falcipa
34	45.8	3.1	82588	21	AAF22301	BAC containing rep
35	45.4	3.1	1082138	21	AAF22305	Arabidopsis thalia
36	45.4	3.1	1335	21	AAA49932	Human delta-5-desa
37	45.4	3.1	1335	22	AAF25236	Nucleotide sequenc
38	45	3.1	451	22	AAH50922	Lipid modification
39	45	3.1	451	22	AAH56909	P patens lipid met
40	45	3.1	451	22	AAH56929	P patens lipid met
41	45	3.1	1465	21	AAAC2244	Arabidopsis thalia
42	44	3.0	347	22	AAF68420	Human lung tumour
43	44	3.0	482	23	AAV74465	DNA encoding novel
44	44	3.0	1478	19	AAV34206	Human secreted pro
45	44	3.0	1700	21	AAZ93706	CYB5RP fatty acid

ALIGNMENTS

RESULT 1

AAAX76589  
ID AAX76589 standard; CDNA; 1463 BP.

XX  
AC AAX76589;

XX  
DT 11-AUG-1999 (first entry)

XX  
DE Caenorhabditis elegans Delta 6 desaturase Ced6.1 encoding CDNA.

XX  
KW Caenorhabditis elegans; C. elegans; Delta 6 desaturase; Ced6.1;

KW gamma-linolenic acid; GLA; plant fatty acid; octadecatetraenoic acid;

KW OTA; eicosanoid; eczema; mastalgia; atherosclerosis; coronary disease;

KW hypercholesterolaemia; diabetic neuropathy; viral infection; acne;

XX  
OS Caenorhabditis elegans.

XX  
FH Key Location/Qualifiers

FT CDS 11..1342

FT FT /\*tag= a

XX  
PN WO9927111-AL.

XX  
PD 03-JUN-1999.

XX  
PF 24-NOV-1998; 98WO-GB03507.

XX  
PR 24-NOV-1997; 97GB-0024783.

XX  
PA (UYBR-) UNIV BRISTOL.

PI Napier JA;

XX WPI; 1999-370905/31.

DR P-PSDB; AAY17751.

XX Desaturase enzymes, the genes encoding them and their uses

XX Claim 29; Fig 1; 44pp; English.

XX The present sequence encodes Caenorhabditis elegans Delta 6 desaturase, designated Cdp6.1. Desaturase enzymes (I) may be used as immunogens to raise and select antibodies (which may be used in immunoassays, and diagnostic tests to detect the presence of (I) in a sample, or to purify (I)) or as a selectable marker for transformation, especially transformations involving plants. (I) can be used to produce gamma-linolenic acid (GLA) (and derivatives of it), which is a high value plant fatty acid that is widely used in medicine for the preparation of compositions for treating disorders associated with deficiencies in GLA or deficiencies in metabolites derived in vivo from GLA, such as octadecatrienoic acid (OTA) and eicosanoids. Disorders that may be treated with GLA and OTA include eczema, mastalgia, atherosclerosis, hypercholesterolaemia, coronary disease, diabetic neuropathy, viral infections, acne, hypertension, cirrhosis and cancer. The nucleotide sequences (II) encoding (I) may be used as probes or primers. Probes may be used to identify and purify nucleic acids and so may be used in diagnosis to detect the presence of (II) in a sample. Primers are useful for amplifying DNA by polymerase chain reaction (PCR). (II) may also be used to prepare an organism that is either chill resistant or that accumulates GLA or metabolites derived from GLA. Hybridizing DNA molecules may be used as anti-sense molecules to alter the expression of (II) by binding to it and preventing transcription. Hybridizing molecules may also be provided as ribozymes which regulate expression by cleaving RNA molecules.

XX Sequence 1463 BP; 445 A; 288 C; 277 G; 453 T; 0 other;

Query Match 99.2%; Score 1451; DB 20; Length 1463;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1462; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 gctccacaaaatggctgcgacagaagatgctccgggttcgaatgaaggctgatggcaa 60  
 Db 1 gctccacaaaatggctgcgacagaagatgctccgggttcgaatgaaggctgatggcaa 60

Qy 61 atggctctaccttagcagggaattggtgaagaacatccaggagagctgttattgaaca 120  
 Db 61 atggctctaccttagcagggaattggtgaagaacatccaggagagctgttattgaaca 120

Qy 121 atatagaattcggatgctactcatattttccacgctttccacgaagatctctcaggc 180  
 Db 121 atatagaattcggatgctactcatattttccacgctttccacgaagatctctcaggc 180

Qy 181 ttataagaacttgaccttgcgaaagacacggagagacgatgaattccttgagaaca 240  
 Db 181 ttataagaacttgaccttgcgaaagacacggagagacgatgaattccttgagaaca 240

Qy 241 attggaagaagacttgcaaaagtgtatcaatgtatcagcatatgatgtcagtggtgc 300  
 Db 241 attggaagaagacttgcaaaagtgtatcaatgtatcagcatatgatgtcagtggtgc 300

Qy 301 acaagaagaagaaaatggttgatcatcttcgaaaaactacgacagaagcttcgatgatgg 360  
 Db 301 acaagaagaagaaaatggttgatcatcttcgaaaaactacgacagaagcttcgatgatgg 360

Qy 361 attaatgaagaacaaatgaacatacttctgtttaaaagcgatttcaaacatttcaattat 420  
 Db 361 attaatgaagaacaaatgaacatacttctgtttaaaagcgatttcaaacatttcaattat 420

Qy 421 ggcatttgcatcttctcagtagtcttggatggttatattctctcagtagtcttatttagc 480  
 Db 421 ggcatttgcatcttctcagtagtcttggatggttatattctctcagtagtcttatttagc 480

RESULT 2

AAAS1232

ID AAAS1232 standard; cDNA; 1461 BP.

XX

Qy 481 acttgcatggcaacaatttcggatggttaacacatgagttctgccaatcaacagccaacaaa 540  
 Db 481 acttgcatggcaacaatttcggatggttaacacatgagttctgccaatcaacagccaacaaa 540

Qy 541 gaacagaccttgaatgatactatttcttctgttcttctgttaattcttacaaggatttcc 600  
 Db 541 gaacagaccttgaatgatactatttcttctgttcttctgttaattcttacaaggatttcc 600

Qy 601 aagagatgggtgggaaggacaagcataaacacatcatcagctgccacaaatgtaattgatca 660  
 Db 601 aagagatgggtgggaaggacaagcataaacacatcatcagctgccacaaatgtaattgatca 660

Qy 661 tgaagtgatcatcagcttggcaccacttttcgcattttatccagagagatttgtgcaagta 720  
 Db 661 tgaagtgatcatcagcttggcaccacttttcgcattttatccagagagatttgtgcaagta 720

Qy 721 taaggccaagctttgaaaaagaacattctcaagatgtgtacatatcaaatctctatttccac 780  
 Db 721 taaggccaagctttgaaaaagaacattctcaagatgtgtacatatcaaatctctatttccac 780

Qy 781 cgcaatgcttccaatgctccgtttctcagtggtcagtcagttcagttcaatgggtatttcaa 840  
 Db 781 cgcaatgcttccaatgctccgtttctcagtggtcagtcagttcagttcaatgggtatttcaa 840

Qy 841 agagaatcaaatggagtacaagggtctatcaagaagaatcattctggggagcaagcaacaat 900  
 Db 841 agagaatcaaatggagtacaagggtctatcaagaagaatcattctggggagcaagcaacaat 900

Qy 901 tttggacattggcttggattctctcaattgttcttattacacacatggcacttcg 960  
 Db 901 tttggacattggcttggattctctcaattgttcttattacacacatggcacttcg 960

Qy 961 ggttgcttatttctattttcacaaatggaggagccttttggatgctcagtagtcac 1020  
 Db 961 ggttgcttatttctattttcacaaatggaggagccttttggatgctcagtagtcac 1020

Qy 1021 tttaacaccataactctgttgataagtagtaccagccaattctcgaaattttaaacacacttcgc 1080  
 Db 1021 tttaacaccataactctgttgataagtagtaccagccaattctcgaaattttaaacacacttcgc 1080

Qy 1081 cgccttcaaatatttgaccacacacacacacacatgactccatctccattcattgattggtcttg 1140  
 Db 1081 cgccttcaaatatttgaccacacacacacacacatgactccatctccattcattgattggtcttg 1140

Qy 1141 ggggtgagctcaattatcagatcgagcaccactgttcccaacaaatgccacgttgcaatct 1200  
 Db 1141 ggggtgagctcaattatcagatcgagcaccactgttcccaacaaatgccacgttgcaatct 1200

Qy 1201 gaatgcttgcgtgaaaataatgtgaaagaatggtgcaaaagagaataatcttcttaccctcgt 1260  
 Db 1201 gaatgcttgcgtgaaaataatgtgaaagaatggtgcaaaagagaataatcttcttaccctcgt 1260

Qy 1261 cgatgactactttgacggatatgcaatgaaattgcaacaaatgaaaaataatggtcgagca 1320  
 Db 1261 cgatgactactttgacggatatgcaatgaaattgcaacaaatgaaaaataatggtcgagca 1320

Qy 1321 cattcaagctaaagctgcctaaacaactcgggtgttcaaaaagtgttttcttctg-tttttt 1379  
 Db 1321 cattcaagctaaagctgcctaaacaactcgggtgttcaaaaagtgttttcttctg-tttttt 1379

Qy 1380 aaatttaattcttgaataattattgttttccgtcaattcttccatcttccatccctttctgg 1439  
 Db 1380 aaatttaattcttgaataattattgttttccgtcaattcttccatcttccatccctttctgg 1439

Qy 1440 agaaataaaacctgttttttcaa 1462  
 Db 1441 agaaataaaacctgttttttcaa 1463





```

Db 1291 caattccgaatatgtgc 1307
||||| ||||| |||
RESULT 4
AAAS1233
ID AAA51233 standard; cDNA; 1275 BP.
XX
AC AAA51233;
XX
DT 26-SEP-2000 (first entry)
XX
DE E. gracilis fatty acid delta-8 desaturase coding sequence.
XX
KW Fatty acid delta-8 desaturase; polyunsaturated; oil; seed;
KW infant formula; dietary supplement; ss.
XX
OS Euglena gracilis.
XX
FH Key Location/Qualifiers
FT CDS 14..1273
FT /tag= a
FT /transl_except= (pos:71..75, aa:ValSerAlaTrp)
FT /product= Fatty_acid_delta-8_desaturase
XX
PN WO200034439-A1.
XX
PD 15-JUN-2000.
XX
PF 06-DEC-1999; 99WO-US28655.
XX
PR 07-DEC-1998; 98US-0111301.
XX
PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI Browse JA, Wallis JG, Watts JL;
XX
DR WPI; 2000-431293/37.
XX
DR P-PSDB; AAY96722.
XX
PT Purified protein having desaturase activity, useful for creating a
PT double-bond between two carbons
XX
PS Claim 8; Fig 7B; 74pp; English.
XX
CC This cDNA encodes Euglena gracilis fatty acid delta-8 desaturase.
CC The sequence is useful for recombinant production of the enzyme and for
CC generating transformed host cells and transgenic plants. The desaturase
CC can be used for creating a double-bond between two carbons. In
CC particular, the enzyme is useful for producing polyunsaturated fatty
CC acids and for generating enzyme-specific antibodies useful for
CC identifying desaturases. Oil-seed plants may be engineered to incorporate
CC the enzyme, so that the plants produce seed oil rich in fatty acids. The
CC fatty acids could be incorporated usefully into infant formula, foods of
CC all kinds, dietary supplements, nutraceutical and pharmaceutical
CC formulations.
XX
SQ Sequence 1275 BP; 294 A; 326 C; 327 G; 328 T; 0 other;

Query Match 7.7%; Score 112.8; DB 21; Length 1275;
Best Local Similarity 48.3%; Pred. NO. 1.6e-17;
Matches 389; Conservative 0; Mismatches 402; Indels 15; Gaps 2;

QY 454 gtattactctgcgatgtttattagcacttgcattggaacaaatcggatggttaacaca 513
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 gtattctattggggcagtggtgtctgggatgcactatcaacagatgggtgctttctca 441
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 514 tgagttctgcatacaacagcaacaagaacagaccttgaatgatactatttttgtt 573
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 tgacattgcaaccacagacttcaagaacccggagactggaaacacccctcggtggactgg 501
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 574 ctttggtaatttcttacaaggattttcaagagattggttggaggacaagcataaactca 633
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 atttggcaatgtctgcgaaggtttttccgtgacatgttggaaaggacagacacaatgcaca 561
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 634 tcacgctgccacaataatgtaattgatcatcagcgttgatcatcagacttgcaccattttcgc 693
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 tcattcggcaaccaatgttcaaggggcacacctgtatttgacaacctcccccccttagc 621
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 694 atttattccaggagatttgcgaagtataaggccagctttgaaaaagcaattctcaagat 753
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 ctggctcgaggatgcgtcacacggcgctcaccgattt-----ccgcgaagct 669
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 754 tttaccataatcaactctctatttccacgcaatggttcccaatgctccgttttcattgcagac 813
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 670 cattcagttccacgacgactatttcttggtcatctgtatttctgttgcgttcatttgggtg 729
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 814 tggtcagtcagttcaatgggtattcaagaagatacaaatggaatgagatcattatcaaaag 873
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 730 tttccagtcgctgttgaccggtgcgagtttgaaggacagagataaccaattctcatcgctc 789
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 874 aaatgcattctgggacgaagcaacaattgttgacatttgggcttg---ggtattctatca 930
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 790 tcagtataagaaggagccatgtggcctcgccctgcactggaccttgaaaggccctgttcca 849
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 931 atgttcttattaccaacatggtccacttcgggttgcgttatttatttatttcaaaaatggg 990
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 850 cttattctttatgccagcatctcactgcgtgttgggtgttttctgttcttcggagctggt 909
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 991 aggaggcctttgattgctcagctagtcactttcaacataaactctgttgataagatcc 1050
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 910 tggcggtctcgccatgtgcgtgttgcattgaaccactccccactggagaagatcgg 969
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1051 agccaattctcgaattttaacaaacttcgcgtctctcctcaaattttgaccacgcgaacat 1110
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 970 ggaccagctcgggatggccatggtattctcgttggccagatccatgagaccatgaacat 1029
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1111 gactccatctccattcattgattggttgggttgggtggaactcaattatcagatcagaccaca 1170
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1030 tcggcgagggtattatcacagattggtttttcggaggcttgaattaccagattgagacca 1089
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1171 cttgtttcccaacaatgcacgcttgaatcgtgaatcgttgcgtgaaatattgtaagaatg 1230
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1090 ttttggcgccctccctccgcacaaacctgcacgctgacagcgttagctaccaggtggaacagct 1149
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1231 gtgcaaaagagaataatcttctctacc 1256
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1150 gtgccagaagacaaacctgccgtatc 1175
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 5
AAZ47129
ID AAZ47129 standard; DNA; 1374 BP.
XX
AC AAZ47129;
XX
DT 28-MAR-2000 (first entry)
XX
DE Fungal delta6-desaturase coding sequence.
XX
KW Polyunsaturated fatty acid; fungus; delta6-desaturase; animal feed;
KW transgenic animal; malnutrition; biosynthesis; ds.
XX
OS Unidentified.
XX
PN WO9961602-A1.
XX
PD 02-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US12088.
XX
PR 29-MAY-1998; 98US-0087578.
XX
PA (OHIS ) UNIV OHIO STATE.
XX

```





CC useful in cosmetic or animal feed formulations. Furthermore, the  
CC compositions may also be used as fat free media or as research reagents.  
XX  
SQ Sequence 1374 BP; 257 A; 386 C; 378 G; 353 T; 0 other;

	Query Match	6.6%	Score 97	DB 22	Length 1374
	Best Local Similarity	47.2%	Pred. No. 1.1e-13		
	Matches 415	Conservative 0	Mismatches 440	Indels 24	Gaps 3
Qy	464	tctcagctttattagcacttgcattggaacaaattcggatggttaaacacatgagttctgc	523		
Db	466	tcggctgcgcttttgggtctgttctggcagcagtcggtggttgggtccacgacttttg	525		
Qy	524	catacaacgccaacaaagacacgaccttgaatgatactattcttcttcttggtaatt	583		
Db	526	catacceaagctctccaggaccgtttctgggtgatcttttcggcgctcttgggaggt	585		
Qy	584	tcttaacaagattttcaagagattgttgggaagacagatacaactcaactcaogctgcc	643		
Db	586	gtctgccagggtctctgctcgttgggtgggaagacaagacaactcaccacccgcc	645		
Qy	644	acaaatgaaattgatcatgacggtgatcagcttggcaccacttttcgacttattcca	703		
Db	646	cccaagctccaacgacgaggaaccgcagattgacaccacccctctgtgacctggagtga	705		
Qy	704	ggagatttgcgaagataaagccagcttttgaaaaagcaat-----tctc	748		
Db	706	catcgctggagatgttctcgagtgctccagatgaggagctgaccgcgcatgtgctcgct	765		
Qy	749	aagattgtacatactaaacatctctatttaccgcgaatgcttccaatgctccgcttctca	808		
Db	766	tctatggtcttgaaaccagacctgttttacttcccattctctggttgcgctctctcc	825		
Qy	809	tggactggtcagtcagttcaatgggtatt-----caagagaatcaaatgagatacaag	862		
Db	826	tgggtccctcagtcacattctcttctgtgcgctaacggtcagggcccacagcccctcgggc	885		
Qy	863	gtctatcaagaagaatgcaattctgggagacaagacaacatttggacattgggtctgggta	922		
Db	886	gcgcgtgtgccactcgtgtgtgcgacagctgcgtctgcgtgcagctgacctgggtac	945		
Qy	923	tctatacaattgtcttattaccacaataggccaact---tcgggttgcttatttcattatt	979		
Db	946	ctcgccacctgtctctgttctatcaaggatcccgtcaacatgctggttacttttgggtg	1005		
Qy	980	tcacaaatgggagagacctttgatgctcacatgactcactttcaaccataactcttt	1039		
Db	1006	tcgcagcggtgtgcggaaacctgttggcgatcgtgtctcgcgtcaaccacaacgggtatg	1065		
Qy	1040	gataagtatccagccaattctcgaattttaaacaacttcgcgctcttccaattttgacc	1099		
Db	1066	cctgtgatctcgaaggaggagcggtgcgatgattgattcttcacgaagcagatcatcag	1125		
Qy	1100	acaagcaaatgactcaactcacttcaattcaattgattggtcttgggtggactcaattatcag	1159		
Db	1126	ggctgctgatgtccaccgggtctatttggcaactggttcacgggttggattgaactatcag	1185		
Qy	1160	atcgagcaccaactgttcccaacaactgccagcttgcgaatcgaatgcttgcgtgaaatat	1219		
Db	1186	atcgagcaccaactgttccctctgatgctcgcgcacactttccaagatccagcctgct	1245		
Qy	1220	gtgaagaagtgtgcaagagaaataactctcttacctcgtctcgatgactactctttgacgga	1279		
Db	1246	gtcgagacctgtgcaaaaagtacaatgtccgataccacacaccacgggtatgatcgaggga	1305		
Qy	1280	tatgcaatgaatttgcaacaattgaaaaatatgctgtgag	1318		
Db	1306	actcgacaggtctttagcgtctctgaacaggtcttccaaq	1344		

RESULT 7  
AAV63624

Query Match	6.6%;	Score 97;	DB 19;	Length 1617;
Best Local Similarity	47.2%;	Pred. No. 1.2e-13;		
Matches 415: Conservative	0;	Mismatches 440;	Indels 24;	Gaps 3;

AAV63624 standard; cDNA; 1617 BP.	AAV63624;	15-FEB-1999 (first entry)	cDNA encoding a delta-6 desaturase enzyme.	Fatty acid; delta-6 desaturase; polyunsaturated fatty acid; malnutrition; inflammation; rheumatoid arthritis; asthma; psoriasis; cancer; diabetes; eczema; platelet aggregation; vasodilation; cholesterol level; endometriosis; premenstrual syndrome; myalgic encephalomyelitis; chronic fatigue; AIDS; multiple sclerosis; acute respiratory syndrome; hypertension; inflammatory skin disorder; ss.
Mortierella alpina.				
Key	Location/Qualifiers			
CDS	71..1444			
FT	/*tag= a			
FT	/product= delta-6 desaturase			
PN	WO9846763-A1.			
XX				
PD	22-OCT-1998.			
XX				
PF	10-APR-1998; 98WO-US071136.			
XX				
PR	11-APR-1997; 97US-0834655.			
XX				
PA	(ABBO ) ABBOTT LAB.			
PA	(CALJ ) CALGENE LLC.			
XX				
PI	Chaudhary S, Huang Y, Knutzon D, Leonard AE, Mukerji P; Thurmond J;			
XX				
DR	WPI; 1998-594582/50.			
DR	P-PSDB; AAW84137.			
XX				
PT	New isolated fatty acid desaturase enzymes - used for the production of polyunsaturated fatty acids for use in, e.g. pharmaceutical compositions, nutritional compositions, cosmetics or animal feed			
PT				
XX	Claim 1; Fig 3A-E; 165pp; English.			
XX				
CC	The present sequence encodes a Mortierella alpina fatty acid delta-6 desaturase enzyme. The enzyme sequence is used in the methods of the invention. The specification describes methods for desaturating a fatty acid and for producing a desaturated fatty acid by expressing increased levels of a desaturase. The present desaturase is an enzyme which introduces a double bond carbons 6 and 7 from the carboxyl end of a fatty acid molecule. The enzyme can be used for desaturating fatty acids. The enzyme can be used to produce polyunsaturated fatty acids, which can be used for treating malnutrition, in pharmaceutical compositions, in cosmetics or in animal feed. The polyunsaturated fatty acids can be used for treating e.g. restenosis after angioplasty, inflammation, rheumatoid arthritis, asthma, psoriasis, cancer, diabetes or eczema or reduce blood pressure. They can also be used to inhibit platelet aggregation, cause vasodilation, lower cholesterol levels, inhibit proliferation of vessel wall smooth muscle and fibrous tissue, reduce or prevent gastro-intestinal bleeding and other side effects caused by non-steroidal anti-inflammatory drugs, prevent or treat endometriosis and premenstrual syndrome, treat myalgic encephalomyelitis and chronic fatigue after viral infections, treat AIDS, multiple sclerosis, acute respiratory syndrome, hypertension and inflammatory skin disorders.			
XX				
Sequence 1617 BP; 310 A; 470 C; 410 G; 427 T; 0 other;				

QY	464	tcTgcattgtttatTtagcaattgcattggccaatTcgcagatggtttaacacatTgattcttcgc	523
Db	536	tcggctgcgccttttgggtctgtctTcTggcagcagTgcggaTggtTggctcagcacttttg	595
QY	524	catcaacagcccaacaagaacagacactTtgaatgatactattTcttTgtttcttTtggaat	583
Db	596	catcaacagTctctccaggaacgctTctcTggggTgatactTctcTggcgccctTctTggaggt	655
QY	584	TctctacaaggattTtccaagagattTgtTgaaggacaagcataaCaactcatcaCgcTgc	643
Db	656	gtctgccaggctTctcTctcTcTgtTgtTgaaggacaagacaCaactcaccacgcgc	715
QY	644	acaaTgtaatTgatcatgacTgTgatctgcactTtggccacactTtTgcatttatTcca	703
Db	716	cccaacTccacTgcgagTatccgacattgacccccctcTgtTgacctggagtTgag	775
QY	704	ggagattTgtcaagtataaaggccagctTtgaaaaagcaat-----tctc	748
Db	776	catTcTctTggagatTctcTcTgatTcccaagTgaggagctTgacccgcatTgTgcTcgt	835
QY	749	aagatTgataccaTacaatctctattTccaCgcaatTgtctccaatTgcctcgtTctcca	808
Db	836	TctatgTctgaaccagacTgTtttactTtccccattctcTgtTtgcCgtctctcc	895
QY	809	TggactTggtcagTcagTtccaT-----caaaagagaaTcaaatTggagTacaag	862
Db	896	TgTgTcTcagTccaattctctTgtcTgcTaaCgTgcaggcccaCaagccctcTggc	955
QY	863	gtctatcaagaagaatCattctTggagacaagcaacaatTgtTgacattTggctTgggta	922
Db	956	gcgcTgtgcCcaattcTgtTgtcTgagcagTgtcTgctTgcTgTgacTgTgacTcTggac	1015
QY	923	TctatcaaatTgtctctTattacaacatTggccact---tcTggtTgtctattTctatt	979
Db	1016	ctgcCcaactTgtctcTgtTcatcaaggatccTgTcaacatTcTgTgTactTttTggTg	1075
QY	980	TcaacaatTgggaggagccTttTgattTgTcTcagTgTcaactTtccaacataaactcTgt	1039
Db	1076	tcgcaggcTgtTgcggaaactTgtTggcagTcTgTctcTgcTcaaccaCacgcTatg	1135
QY	1040	gataagTatccagccaattctcTgaattTtaacaactTgcgcTctTccaattTgacc	1099
Db	1136	ctcTgTatctcaaggaggcTgcatatTgTgattTcttccaagacagatcaTcaag	1195
QY	1100	acacTcaacatTgactccatctccatctcattTgattgTgctTtggTgTgactcaattatca	1159
Db	1196	gTcTgTgTgcacccTgtctctcTcTgTgTgcTgcTgcTgcTgcTgcTgcTgcTgcTgcTgc	1255
QY	1160	atcTgacacacTgtTtccaacaatTgcacTgtTgcaatTgcaatTgctTgctgTgaaatat	1219
Db	1256	atcTgacacacTgtTtccctcTcTgTgTgcTgcTgcTgcTgcTgcTgcTgcTgcTgcTgc	1315
QY	1220	gtgaaagaatTgtTgcaagaagaataatcttccattcTcTcTgcTgactactTtgacgga	1279
Db	1316	gtcTgacccTgtgcaaaaagTacaatTcTcgataccaacacacccTgtatgatcTgagga	1375
QY	1280	tatgcaatTgaattTgcaacaattTgaaaaatTggtcTgag	1318
Db	1376	actTgcagagTcttttagcctcTctgaacTgagTctccaag	1414

## RESULT

AXX00889

AA000889  
ID AA000889 standard: DNA: 1617 BP.

AA  
AC  
AAX00889:

AC  
XX  
DT

**DE** Mortierella alpina delta 6 desaturase encoding DNA.

Delta 6 desaturase; recombinant; fatty acid desaturase; FAD; PUFA; oil; *Neisseria arifina* and desaturase encoding DNA.

polyunsaturated fatty acid; linoleic acid; arachidonic acid; linolenic; stearidonic acid; eicosapentaenoic acid; malnutrition; feeding formula; dietary supplement; prostaglandin; restenosis; angiotlasty; inflammation; rheumatoid arthritis; psoriasis; osteoporosis; cancer; eczema; AIDS; diabetes; cosmetic; animal feed; ss.

*Mortierella alpina*.

Key	Location/Qualifiers
1	1.1
2	2.1
3	3.1
4	4.1
5	5.1
6	6.1
7	7.1
8	8.1
9	9.1
10	10.1
11	11.1
12	12.1
13	13.1
14	14.1
15	15.1
16	16.1
17	17.1
18	18.1
19	19.1
20	20.1
21	21.1
22	22.1
23	23.1
24	24.1
25	25.1
26	26.1
27	27.1
28	28.1
29	29.1
30	30.1
31	31.1
32	32.1
33	33.1
34	34.1
35	35.1
36	36.1
37	37.1
38	38.1
39	39.1
40	40.1
41	41.1
42	42.1
43	43.1
44	44.1
45	45.1
46	46.1
47	47.1
48	48.1
49	49.1
50	50.1
51	51.1
52	52.1
53	53.1
54	54.1
55	55.1
56	56.1
57	57.1
58	58.1
59	59.1
60	60.1
61	61.1
62	62.1
63	63.1
64	64.1
65	65.1
66	66.1
67	67.1
68	68.1
69	69.1
70	70.1
71	71.1
72	72.1
73	73.1
74	74.1
75	75.1
76	76.1
77	77.1
78	78.1
79	79.1
80	80.1
81	81.1
82	82.1
83	83.1
84	84.1
85	85.1
86	86.1
87	87.1
88	88.1
89	89.1
90	90.1
91	91.1
92	92.1
93	93.1
94	94.1
95	95.1
96	96.1
97	97.1
98	98.1
99	99.1
100	100.1

key	LOCATION,
CDS	71...1443

$$/\ast\text{taq} =$$

```

/product= "delta 6 desaturase"

```

WO9846764-A1.

22-OCT-1998

10-APR-1998. 98WQ-IIS07A21

34-00000-1007. 07110-0050095

24-OCT-1997; 97US-0936983.  
11-APR-1997. 97US-0833610

11 - AFK-1997, 3705-0833010.  
11 - APR-1997, 97US-0834033

II APR-1997; 97US-0834655;  
11-APR-1997; 97US-0834655;

(ABBO ) ABBOTT LAB.

(CALJ ) CALGENE LLC.

Chaudhary S. Huang Y. Knutson D. Leonard AE. Mukerji P:

Thurmond J;

WPT: 1999-080739/07

P-PSDB; AAW95504

Nucleic acid construct able to express fatty acid desaturase in plants - useful in human or animal nutrition, as cosmetics and therapeutically, e.g. for restenosis, cancer and diabetes

Claim 1: Fig 3A-E: 210pp: English:

This DNA encodes a Mortierella alpina delta 6 desaturase. The invention relates to a nucleic acid construct that contains at least one of the nucleotide sequences (AAX00889 to AAX00891) encoding M. alpina delta 6, delta 12 and delta 5 desaturases (AAW95504 to AAW95506) respectively, coupled to an expression control sequence functional in plants.

Recombinant plant cells containing at least one DNA encoding a M. alpina fatty acid desaturase (FAD), can be used for the production of polyunsaturated fatty acid (PUFA). These recombinant cells or plants containing them are used to produce oils such as linoleic acid,

arachidonic acid, gamma-linolenic acid, dihomo-gamma-linolenic acid, stearidonic acid and eicosapentaenoic acid (EPA). These plant oils are used: (i) to treat malnutrition; (ii) in infant feeding formulas, or dietary supplements or substitutes, for use in humans or animals; (iii) for treating disorders associated with inadequate consumption or production of PUFA (or their metabolites such as prostaglandins), e.g.

restenosis after angioplasty, inflammation, AIDS, rheumatoid arthritis, psoriasis, osteoporosis, cancer, eczema, diabetes; (iv) as cosmetics, and (v) as animal feeds. Fragments of the DNA are used as probes to isolate related coding sequences. Recombinant plants can produce high yields of PUFA, since new pathways can be created and unwanted ones suppressed.

born; since new patents can be created and awarded ones suppressed, plants can be engineered to express oils of particular PUFA composition, e.g. one similar to that in human milk, and product recovery is simpler than with e.g. fish.

Sequence 1617 BP: 310 A; 470 C; 410 G; 427 T; 0 other;

Query Match

Query Match

Query Match 6.6%; Score 97; DB 20;  
Best Local Similarity 47.2%; Pred. No. 1.2e-13;

Oy 464 tctgcatagtttattagcacttgcattggcaacaattcggatggttaacacatgagttctgc 523

Accession	Gene	Accession	Gene
Q7	403	536	D6



```
Db 776 catcgcttgagagatgtcttcgagatgtccagatgaggagcagaccgcgcatggtgcgct 835
Qy 749 aagattgtaccatacaacatctctattccaccgaagtcttccaaatgctccgtttctca 808
Db 836 ttcatggtctgaacacagacctggttttacttcccactctctctgttgcgcgtctctcc 895
Qy 809 tggactggtcagtcagttcaatgggtatt-----caagagaatcaaatggagtacaag 862
Db 896 tgggtctccagtcacatctctgtgtgctgaacggtcgagcccaacagccctcgaggc 955
Qy 863 gtctatcaaaagaatgcatcttgggagcaagaacaatgtttggacattgggctgggta 922
Db 956 ggcgctgtgccatctcgtttgtcgagcagctgtcgtgtcgatgcactggacctgggtac 1015
Qy 923 ttctatcaatgtctctattaccacatggccact---tcgggtgtctatttcaattatt 979
Db 1016 ctgcgccaccatgtctctgttcatcaagatcccgctcaacatgctggttacttttgg 1075
Qy 980 tcacaaatggaggagcgttttattgtctcacgtagtcactttcaaccataaactctgtt 1039
Db 1076 tcgcaggcgtgtgcggaactgttggcgatcgtgttctcgctcaacacacacaggtatg 1135
Qy 1040 gataagtatccagccaatctctogaatttaacaaacttcgcgcctcttcaaatttgacc 1099
Db 1136 cctgtgatctgaaggaggcgtcgatgattgattcttcacgaagcagatcatcacg 1195
Qy 1100 acagcaacatgactccatctccattcattgattgctgctggttgggtggactcaattatcag 1159
Db 1196 ggtcgtgatgtccacccgggtctatttccaaactggttcacgggtggtgaactatcag 1255
Qy 1160 atcgagcaccactgttcccacaacatgccaactgtgcaatctgaaatgctgctgaaat 1219
Db 1256 atcgagcaccactgttcccttcgactgctcgccaacttttcaaatccagcctgct 1315
Qy 1220 gtgaagaatggtgcgaagagaataatcttcttaccctcgtcgtgatctatttgacgga 1279
Db 1316 gtcgagaccctgtgcaaaagtacaaatgctcgtatccgataccaccaccggtgatogagga 1375
Qy 1280 tatgcaatgaatttgcaacaattgaaataatgctgag 1318
Db 1376 actgcagagggtcttagcgtctgaacgaggtctccaag 1414

RESULT 10
AAF26040
ID AAF26040 standard; CDNA; 2012 BP.
XX
AC AAF26040;
XX
DT 23-APR-2001 (first entry)
XX
DE P. patens delta6-desaturase cDNA.
XX
KW Delta6-desaturase; unsaturated fatty acid; transgenic; oil; lipid;
KW fatty acid; human nutrition; animal nutrition; cosmetic; pharmaceutical;
KW agricultural chemical; ds.
XX
OS Physcomitrella patens.
XX
FH Key Location/Qualifiers
CDS 319..1896
FT /*tag= a
FT /product= "delta6-desaturase"
XX
FN WO200102591-A1.
XX
PD 11-JAN-2001.
XX
PF 04-JUL-2000; 2000WO-EP06223.
XX
PR 06-JUL-1999; 99US-0347531.
PR 30-JUN-2000; 2000DE-1030976.
```

```
XX
PA (BADI ) BASF AG.
XX
PI Heinz E, Girke T, Scheffler J, Da Costa Silva EO;
XX
DR WPI; 2001-123117/13.
DR P-PSDB; AAB46810.
XX
PT Production of unsaturated fatty acids, useful e.g. in nutrition,
PT cosmetics or pharmaceuticals, in organisms transformed with
PT Physcomitrella patens delta-6-desaturase nucleic acid -
XX
PS Claim 1a; Page 38-41; 49pp; German.
XX
CC This invention describes a novel preparation of unsaturated fatty acids
CC (I) by introducing into an organism at least one isolated nucleic acid
CC (II) that encodes a polypeptide (III) with delta6-desaturase activity.
CC Organisms that contain at least 1 wt.% (I), on total fatty acid content,
CC are then selected. (II) is selected from: (a) a 2012 bp sequence (S1),
CC defined in the specification, or its equivalents within the degeneracy of
CC the genetic code; or (b) derivatives of the sequence of (a) that encode a
CC 525 amino acid polypeptide (S2), defined in the specification, or a
CC polypeptide with at least 50% homology with (S2) and practically the same
CC enzymatic activity. The invention also describes (1) transgenic organisms
CC that contain (II); and (2) oils, lipids and fatty acids produced by the
CC new method. The oils, lipids and fatty acids produced by the transformed
CC organisms are used in human or animal nutrition, cosmetics,
CC pharmaceuticals and agricultural chemicals. (III) can also be used, in
CC vitro, for increasing the (I) content of triglycerides. The transgenic
CC organisms have increased contents of (I), or of (I)-containing
CC triglycerides, particularly of gamma-linolenic acid.
XX
SQ Sequence 2012 BP; 482 A; 427 C; 553 G; 550 T; 0 other;

Query Match 5.2%; Score 76.2; DB 22; Length 2012;
Best Local Similarity 56.6%; Pred. NO. 1.3e-08;
Matches 141; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 412 ttcaattatggcattgtcttattcttcagtatcttggatgtattacttctgcattg 471
Db 978 tgcgacgcatgcaataatattgttgagcaagactatttcagcgttttggcttcagcttg 1037
Qy 472 ttattagcacttgcattgcacaaattcggatggttaacacatgattctgcacataca 531
Db 1038 tatgatggctctgtgtttccaaacagtcggatgctatccccatgattttccacaatca 1097
Qy 532 gccacaagaacagacaccttgaatgatactatttcttcttcttggtaatttttaca 591
Db 1098 ggtgttgagacacgctggtcttaataagttgtcgggtatgtgatcggaacgcgcttct 1157
Qy 592 aggattttcaagagattgttggaaggacaagcacaataacactcatcagctgcacaaatgt 651
Db 1158 ggggttttagtcacaggtgtggaaggagaagcacaataaccttcacatcgtctgccaaatga 1217
Qy 652 aattgatca 660
Db 1218 atgcgatca 1226

RESULT 11
AAV34398
ID AAV34398 standard; DNA; 1684 BP.
XX
AC AAV34398;
XX
DT 02-MAR-1999 (first entry)
XX
DE Borage delta-6 desaturase gene.
XX
KW Upstream region; regulatory region; sunflower; albumin; seed; expression;
KW lipid metabolism; delta-6 desaturase; transgenic plant; ds.
XX
```



PT particularly to increase content of gamma-linolenic acid

XX  
PS Example 2; Page 59-61; 101pp; English.

The present sequence encodes borage delta-6 desaturase (see AA098130), an enzyme that catalyses the conversion of linoleic acid to gamma-linolenic acid (GLA). Delta-6 desaturase cDNA was isolated from a borage membrane-bound polysomal cDNA library using a partial clone, obtained from an EST database search, as probe. The borage delta-6 desaturase nucleic acid can be operably linked to the seed-specific 5' regulatory region (see AA24916) of the *Arabidopsis thaliana* oleosin AtS21 gene in claimed expression cassettes of the invention. Transgenic plants, e.g. sunflower, soybean, maize, tobacco, cotton, peanut, oilseed rape or *Arabidopsis* are obtained that show increased levels of GLA or octadecatetraenoic acid. The levels of desirable fatty acids in oilseed crops can be manipulated to provide seed oils of use in human health and industrial applications.

Sequence 1684 BP; 430 A; 277 C; 358 G; 619 T; 0 other;

Query Match 4.7%; Score 68.4; DB 20; Length 1684;  
Best Local Similarity 44.2%; Pred. No. 9.5e-07;  
Matches 423; Concentration 0.0; Mismatch 526; Total 0

Qy	381	cataattctcgtttaagcgattcaacactttcaattatggcatttgcatatttatcttc	440
Db	386	caacttttgcctttatagaatgctgtttgctatgagtgttatagggttttgtttgtg	445
Qy	441	agtatcttggatggatatatactcttcgatgtttatttagcacttgcattgccacaattcg	500
Db	446	agggtgttttggtaeaatttgttttcgggtgtttgatgggtttcttttggattcagagtg	505
Qy	501	gatvgtttaacacatgagttctgcctaacaccgaacaaagaaacagacctttgaatgata	560
Db	506	gttggattggacatgatctgggcattatagttagtcttgattcaagggttaataagt	565
Qy	561	ctattctttgtctcttggtaattcttcaaggattttcaagagatttccaaagagattgttgggaagaca	620
Db	566	ttaagggtattttgtctgaattgtcttccaagtaaatgaattatgtgttggcggaaaTgga	625
Qy	621	agcataacactcatcgcgtgcacaaatgttaattgatcatgacgggtgatatacgacttgg	680
Db	626	accataatgacatacacattgctgttaatagccttgaataTgaccttgatttacaata	685
Qy	681	caccaatttcgcatttatccaggagatttgtcaagtataaaggccaagctttgaaaaag	740
Db	686	taccattctctgtgtgtctccaagtttttggttcactcacctctoaatttctatgaga	745
Qy	741	caattctcaagattgtaccataTcaactctctatttccacgcgaatgcttc--caatgc	797
Db	746	aagggttgacttttgacttttatcaagattctttgttaagtattcaacattggacattt	805
Qy	798	tcggtttctcattggaactgtcagtcagttccaalggttatccaagagaatacaaatggagt	857
Db	806	acctattatgtgtcgtctagctccaatgtatgtcaactctctcataatgttgttga	865
Qy	858	acaaggtctcatcaaagaaatgatctctggagcaagcaacaatttgttggacattggcctt	917
Db	866	ccaagagaaatgtgtcctaTcagctcaggaactcttgggtgctcagtggtctcgattt	925
Qy	918	gggtattctataaattgtctcttattaccaacatggccaactcgggttgccttatttoatta	977
Db	926	ggtacccggttgcgtgtttctcttTtgcctaattgggttggaaagaattatgtttgtattg	985
Qy	978	tttcaacaaatgggaggagcccttttggatgtctcagctagtcacttttceaacataactctg	1037
Db	986	caagtttatcagtgactggaaTgcacaagtcca--gtctctcttgaaccacttctctt	1042
Qy	1038	ttgataagtatccagccaattctcgaatttttaacaaactctgcgcgtctcttcaaatTTga	1097
Db	1043	caagttgtttatgttgaagacct---aaagqaaTaattgttttTgaacaaacacagctatg	1099

Query Match		4.7%;	Score 68.4;	DB 17;	Length 1685;
Best Local Similarity		44.2%;	Pred. No. 9.5e-07;		
Matches 423;		Conservative	0;	Mismatches 526;	Indels 9; Gaps 3;
QY	381	catattctctgtttaaagcgatttcaacactttcaattatgcatgttgcatatttatcttc	440		
Db	387	caactttgtctttatagcaatgctgttgctatgagtttatggggtttgtttgtg	446		
QY	441	agtatcttggtatattactctgtctgtttatagcaacttgcatgcaacaattcg	500		
Db	447	aggggtttgtgtacattttgttctgggtgttgatgggtttctttggttgcagagt	506		
QY	501	gatggttaacacatgagttctgocatacaagccaaagcaacagacactttgaaatgata	560		
Db	507	gttgattggacatgatgctggcattatattgtagtctgtattcaaggcttaataagt	566		
QY	561	ctattttctgtttgttaattttcacaaggattttcaagagattttcaagagattggtggaagaca	620		
Db	567	ttatgggtatttttgcgcacaaattgtcttcaggaataagattggtgtggaatgga	626		
QY	621	agcataaactcatcacgctgcacaaatgtaattgatcatgacggtgatctgacttgg	680		
Db	627	accataatgcacatcacattgctgttaataagcttgaaatgacccctgtttacaatata	686		
QY	681	caccacttttcgcaatttctcaggagatttgcgaagtataaggccagctttgaaaaag	740		
Db	687	taccattcttgtgtctccaaatttttgggttcactcacctctcattttctatgaga	746		
QY	741	caattctcaagatttacatatcaacatctctattttaccgcaatgcttc---caatgc	797		
Db	747	aaaggttgacttttactctttatcaagattctttgtaagttatcaacattggacattt	806		
QY	798	tcggtttctcagctgcagtcagttcaatgggtatttcaaaagagaatcaaatggagt	857		
Db	807	accctattatgtctgtaggtcacaatgatgtatgatacaactctctcataaagtgttga	866		
QY	858	acaaggtctatcaagaataatgattctgtgggacaagcaacaattgttgacattgggctt	917		
Db	867	ccaagagaaatgtctctatgagctcaggaaactctgtggatgcctagtgtctcgattt	926		
QY	918	gggtattctcaattgtttctattaccaaatgcccacttcgggttgcttatttcaata	977		
Db	927	ggatcccgctgtctgtttctgttgcctaatggggtgaaagaattatgttgtttatg	986		
QY	978	tttcaaaatgggaggagccttttgattgctcaogtagtcttcaaccataaactctg	1037		
Db	987	caagtttatcagtgactggaatgcacaagtcca---gttctcttgaaccactctctt	1043		
QY	1038	ttgataagttaccagocaaattctcgaattttaacaaacttcgccgtcttcaaatttga	1097		
Db	1044	caagtgtttatgttgaaagcct---aaagggaaataattggtttgagaaacaacggatg	1100		
QY	1098	ccacagcaacatgactccatctccattcattgattgcttgggttggaactcaattatc	1157		
Db	1101	ggacatttgacattctgtctctctgtgattggttcttcatggttgattgcaattcc	1160		
QY	1158	agatgagcacacactgtttcccaacaatgcccggttgcaatctgaatgcttgcgtgaaat	1217		
Db	1161	aaatgagcatcatttttcccaagatccttagatgcaaccttaggaaactctgcctc	1220		
QY	1218	atgtgaagaaatggttgcgaagaaataatcttcttaccctctgcctgactacttgacy	1277		
Db	1221	acgtgatcgagttagcaagaacaataatttgccttacaattatgcatctttctccaagg	1280		
QY	1278	gatatgcaatgatttgcaacaattgaaaaataatggctgagcacattcaagctaaagc	1335		
Db	1281	ccaatgaaatgacatcgaacattgagggaacacagcattgcaggctaggatataac	1338		

RESULT 14

AAF25730

ID AAF25730 standard; DNA; 1467 BP.

XX

AC

XX

DT

XX

DE

XX

KW

KW

KW

XX

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

AAF25730;

06-APR-2001 (first entry)

C. purpureus delta6-acetylenase/delta6-desaturase DNA SEQ ID NO 3.

Delta6-acetylenase; delta6-desaturase; unsaturated fatty acid;

transgenic plant; plant oil; triglyceride; nutrition; animal feed;

cosmetic; ds.

Ceratodon purpureus.

Key Location/Qualifiers

CDS 10..1461

/\*tag= a

/product= "delta6-acetylenase/delta6-desaturase"

WO200075341-A1.

14-DEC-2000.

07-JUN-2000; 2000WO-EP05274.

07-JUN-1999; 99DE-1025718.

22-DEC-1999; 99DE-1062409.

(BADI ) BASF AG.

Heinz E, Stymne S, Lee M, Girke T, Sperling P, Zaehringer U;

WPI; 2001-112150/12.

P-PSDB; AAB46436.

Nucleic acid encoding delta6-acetylenase or desaturase, useful for

producing plant oils with increased content of unsaturated fatty acids

Claim 1a; Page 46-48; 69pp; German.

This invention describes a novel isolated nucleic acid (I) encoding polypeptides (II) with Delta6-acetylenase and/or Delta6-desaturase activity. The invention also describes (a) amino acid sequences encoded by (I); (b) an expression cassette (EC) containing (I) linked to one or more regulatory sequences; (c) a vector containing (I) and EC; (d) organisms containing (I), EC or the vectors of (c); (e) preparation of unsaturated fatty acids (A) or triglycerides (TG) with increased content of (A) by introducing (I) or EC into an oil-producing organism; (f) proteins (IIa) of 172 aa or 178 aa (given in the specification); (g) production of (A) or TG by using (Ia); and (h) (A) and TG produced by method (g). (I) are used to produce transgenic plants (or other organisms) that produce oils or triglycerides (TG) with increased content of unsaturated fatty acids (A) and to isolate related sequences by homology screening. (A), or TG containing them, are useful in human nutrition (e.g. infant foods), animal feeds, pharmaceuticals and cosmetics. Derivatives of (I), e.g. antisense sequences or ribozymes, can be used to suppress expression of (II), resulting in oils with increased content of saturated fatty acids.

Sequence 1467 BP; 356 A; 349 C; 379 G; 383 T; 0 other;

Query Match 4.2%; Score 62; DB 22; Length 1467;

Best Local Similarity 44.0%; Pred. No. 3.2e-05;

Matches 376; Conservative 0; Mismatches 460; Indels 18; Gaps 2;

QY 464 tctgcattttattagcacttgcatggcacaattcggttggttaacacatgattctgc 523

II III III III III III III III III III III III III III III III

Db 595 tcagccagtttgatgggtctcttcgccacagtggtggtgcttgcctatgattccct 654

II III III III III III III III III III III III III III III III

QY 524 catcaacagcccaacaaagacagaccttgatgatactattctttgtttgttgtaat 583

II III III III III III III III III III III III III III III III

Db 655 catcaacaggttctttgagaacccgtaccggaactcttcttggctatttttcggcaat 714

II III III III III III III III III III III III III III III III









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2002, 16:04:11 ; Search time 31.81 Seconds  
(without alignments)  
1546.862 Million cell updates/sec

Title: US-09-555-093-2  
Perfect score: 2410  
Sequence: 1 MVVDKNASGLRMKVGKWL.....YAMNLOOLKNMAEHIOAKAA 443

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2410	100.0	443	20 AAY17751	Caenorhabditis ele
2	2382	98.8	473	21 AAY51353	Protein b5cae with
3	1092.5	45.3	447	20 AAY21891	C. elegans delta 5
4	1092.5	45.3	447	21 AAY96721	C. elegans fatty a
5	1079	44.8	454	21 AAY51355	Protein 25bce with
6	597.5	24.8	422	21 AAY96722	E. gracilis fatty
7	592	24.6	520	22 AAB46440	C. purpureus delta
8	590.5	24.5	525	21 AAY51354	Protein b5pp with
9	590.5	24.5	525	22 AAB46810	P. patens delta6-d
10	559	23.2	483	22 AAB46435	C. purpureus delta
11	559	23.2	483	22 AAB46436	C. purpureus delta

12	557	23.1	457	19 AAW84137	A delta-6 desatura
13	557	23.1	457	20 AAW95504	Mortierella alpina
14	557	23.1	457	21 AAY92599	M. alpina delta-6
15	557	23.1	457	21 AAY56045	Fungal delta6-des
16	557	23.1	457	22 AAB31684	Amino acid sequenc
17	555	23.0	457	20 AAW85121	A delta-6 desatura
18	501	20.8	355	19 AAW84139	Desaturase enzyme
19	463.5	19.2	449	21 AAG29290	Arabidopsis thalia
20	462.5	19.2	462	21 AAY71552	Corn sphingolipid
21	460.5	19.1	458	21 AAY51348	Sphingolipid desat
22	460	19.1	448	21 AAY71551	Florida bitterbush
23	456.5	18.9	446	20 AAW85122	A delta-6 desatura
24	454.5	18.9	448	17 AAR98455	Borage delta-6-des
25	454.5	18.9	448	19 AAW67471	Borage delta-6 des
26	454.5	18.9	448	20 AAW98130	Sunflower HADES pr
27	454.5	18.9	448	21 AAY51349	Wheat sphingolipid
28	450.5	18.7	469	21 AAY71555	Sheaf sphingolipid
29	443.5	18.4	449	21 AAY51333	B. napus sld1 prot
30	434	18.0	449	21 AAG07392	Arabidopsis thalia
31	434	18.0	517	21 AAG07391	Arabidopsis thalia
32	433	18.0	449	21 AAG53861	Arabidopsis thalia
33	433	18.0	449	21 AAY51334	A. thaliana sld1 p
34	418	17.3	450	21 AAY71554	Soybean sphingolip
35	412	17.1	87	19 AAW84144	Desaturase enzyme
36	411	17.1	444	21 AAY95445	Human delta-5-des
37	411	17.1	444	22 AAB31686	Amino acid sequenc
38	411	17.1	501	22 AAM93314	Human polypeptide,
39	410.5	17.0	326	21 AAG29291	Arabidopsis thalia
40	406	16.8	425	22 AAM25725	Human protein sequ
41	406	16.8	444	21 AAY97538	Human fatty acid d
42	406	16.8	444	22 AAB94041	Human protein sequ
43	398	16.5	432	21 AAY95446	Human delta-5-des
44	398	16.5	746	19 AAW84156	Human desaturase e
45	398	16.5	746	20 AAW85135	A desaturase enzym

ALIGNMENTS

RESULT 1  
AAY17751  
ID AAY17751 standard; Protein; 443 AA.  
XX  
AC AAY17751;  
XX  
DT 11-AUG-1999 (first entry)  
XX  
DE Caenorhabditis elegans Delta 6 desaturase Ced6.1.

XX  
KW Caenorhabditis elegans; C. elegans; Delta 6 desaturase; Ced6.1;  
KW gamma-linolenic acid; GLA; plant fatty acid; octadecatetraenoic acid;  
KW OTA; eicosanoid; eczema; mastalgia; atherosclerosis; coronary disease;  
KW hypercholesterolaemia; diabetic neuropathy; viral infection; acne;  
KW hypertension; cirrhosis; cancer.

XX  
OS Caenorhabditis elegans.

XX  
PN WO9927111-A1.

XX  
PD 03-JUN-1999.

XX  
PF 24-NOV-1998; 98WO-GB03507.

XX  
PR 24-NOV-1997; 97GB-0024783.

XX  
PA (UYBR-) UNIV BRISTOL.

XX  
PI Napier JA;

XX  
DR WPI; 1999-370905/31.

XX  
DR N-PSDB; AAX76589.

XX  
PT Desaturase enzymes, the genes encoding them and their uses

XX PS Claim 1; Fig 1; 44pp; English.

XX CC The present sequence is *Caenorhabditis elegans* Delta 6 desaturase,

CC designated Cgd6.1. Desaturase enzymes (I) may be used as immunogens to

CC raise and select antibodies (which may be used in immunoassays, and

CC diagnostic tests to detect the presence of (I) in a sample, or to purify

CC (II) or as a selectable marker for transformation, especially

CC transformations involving plants. (I) can be used to produce gamma-

CC linolenic acid (GLA) (and derivatives of it), which is a high value

CC plant fatty acid that is widely used in medicine for the preparation of

CC compositions for treating disorders associated with deficiencies in GLA

CC or deficiencies in metabolites derived *in vivo* from GLA, such as

CC octadecatrienoic acid (OTA) and eicosanoids. Disorders that may be

CC treated with GLA and OTA include eczema, mastalgia, atherosclerosis,

CC hypercholesterolaemia, coronary disease, diabetic neuropathy, viral

CC infections, acne, hypertension, cirrhosis and cancer. The nucleotide

CC sequences (II) encoding (I) may be used as probes or primers. Probes may

CC be used to identify and purify nucleic acids and so may be used in

CC diagnosis to detect the presence of (II) in a sample. Primers are useful

CC for amplifying DNA by polymerase chain reaction (PCR). (II) may also be

CC used to prepare an organism that is either chill resistant or that

CC accumulates GLA or metabolites derived from GLA. Hybridizing DNA

CC molecules may be used as anti-sense molecules to alter the expression of

CC (II) by binding to it and preventing transcription. Hybridizing

CC molecules may also be provided as ribozymes which regulate expression by

CC cleaving RNA molecules.

XX CC Sequence 443 AA;

Query Match 100.0%; Score 2410; DB 20; Length 443;  
Best Local Similarity 100.0%; Pred. No. 1.7e-229;  
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDKNASGLRMKVDGKWLILSEELVKHHPGGAVIEQYRNSDATHIFHAFHEGSSQAYKQ 60  
Db 1 mvdknasglrmkvdgkwlyiseelvkhhpggavieqyrnsdathifhafhegssqaykq 60

Qy 61 LDLLKKHGEHDEFLEKQLEKRLDKVDINVSAYDVSVAQEKMWESFEKRLKLDHDDGLMK 120  
Db 61 ldllkkhgeheflekqlekrldkvdlnvsaydvsavagkkmvsefekrlklhdgdgllmk 120

Qy 121 ANETYFLFKALSTLSIMAFYLYGLWYITTSACLLALAWQOFGWLTHFCHQOQTKNRP 180  
Db 121 anetyflfkalstlsimafylyglwyitsacllalawqgfwlthefchqgqtknrp 180

Qy 181 LNDTISLFFGNFLOGFSRDWKKDKHNTHTHAATNVLDHGDIDDLAPLFAFIPGDLCKYKAS 240  
Db 181 lndtislffgnlfgfsrdwkkdkhnthtaantnvldhgdidldlaplfafipgdickykas 240

Qy 241 FEKATLKIVPYOHLVFTAMLPLRFSTGQSVQWVFKENOMEKYKYORNAFEGQATIVGH 300  
Db 241 fekalikivpyqhlvftamlplrfstgqsvqvwvfkengmeykyvqrnatwedqativgh 300

Qy 301 WAWFYQLFLPTWPLRVAYFIISOMGGGLTIAHVVTENHNNSVDKYPANSRLNFAALQ 360  
Db 301 wawfyqlflptwplrvayfiisomgggliahvvtfnhnsvdkypansrlnnfaalq 360

Qy 361 ILTRNMTSPSPIDMLWGLNLYQIEHHUFLPTMPRCNLNACVKYKWKCKENNLPLYVDY 420  
Db 361 llttrnmtspsfidmlwglnglyqiehhflptmprcnlnacvkyykewckennlplyvdy 420

Qy 421 FDGYAMNIQQLKMAEHIOAKAA 443  
Db 421 fdgyamniqqikmaehioakaa 443

RESULT 2  
AA51353  
ID AA51353 standard; Protein; 473 AA.  
XX  
AC AA51353;

XX DT 27-APR-2000 (first entry)

XX DE Protein b5cae with delta6 fatty acid desaturase activity.

XX KW Sphingolipid desaturase; sld1; sphingobase; ceramide; capnoid;

XX KW transgenic plant; crop plant; delta-8-unsaturated long-chain base;

XX KW tolerance; resistance; soil salinity; ion stress; toxicity; drought;

XX KW cold; frost; phytopathogenic microorganism; flowering time; cosmetic;

XX KW pharmaceutical; food; chemical raw material.

XX OS Unidentified.

XX PN DE19828850-A1.

XX PD 30-DEC-1999.

XX PF 27-JUN-1998; 98DE-1028850.

XX PR 27-JUN-1998; 98DE-1028850.

XX PA (GVSE-) GVS GES ERWERB & VERW LANDWIRTSCHAFTLICH.

XX PI Heinz E, Zaehrer U, Schmidt H, Sperling P;

XX DR WPI; 2000-127549/12.

XX PT New sphingolipid desaturase that selectively introduces double bond

XX PS into sphingolipids and capnoids -

XX PS Disclosure; Page 38-39; 62pp; German.

XX CC This invention describes a novel sphingolipid desaturase that selectively

CC introduces a double bond into the sphingobase of the ceramide residue of

CC sphingolipids and capnoids. A DNA sequence encoding the sphingolipid

CC desaturase, or a vector containing the DNA sequence, can be used to

CC produce transgenic plants, especially crop plants, with an increased or

CC decreased delta-8-unsaturated long-chain base content or an altered

CC delta-8-unsaturated long-chain base cis/trans ratio, especially to

CC compensate for a delta-8-unsaturated long-chain base deficiency, to

CC exclude production of delta-8-unsaturated bases, to increase tolerance

CC or resistance to soil salinity, ion stress or toxicity, drought, wet

CC conditions, cold or frost and/or phytopathogenic microorganisms, or to

CC alter size growth and flowering time. Cells, transgenic organisms or

CC plants containing the DNA sequence can be used to produce sphingolipids

CC and capnoids with unsaturated sphingobases. The sphingolipids or capnoids

CC can be used in cosmetics, pharmaceuticals and foods and as chemical raw

CC materials. This sequence represents a protein which has delta6 fatty acid

CC desaturase activity which is described in the method of the invention.

XX SQ Sequence 473 AA;

Query Match 98.8%; Score 2382; DB 21; Length 473;  
Best Local Similarity 93.4%; Pred. No. 1.1e-226;  
Matches 442; Conservative 1; Mismatches 0; Indels 30; Gaps 1;

Qy 1 MVDKNASGLRMKVDGKWLILSEELVKHHPGGAVIEQ----- 37  
Db 1 mvdknasglrmkvdgkwlyiseelvkhhpggavieqysiplnknletrglttrgsn 60

Qy 38 -----YRNSDATHIFHAFHEGSSQAYKQLDLLKKHGEHDEFLEKQLEKRLDKVDINVS 90  
Db 61 aldillyfyrnsdathifhafhegssqaykqldllkkhgeheflekqlekrldkdvins 120

Qy 91 AYDVSVAQEKMWESFEKRLKLDHDDGLMKANETYFLFKALSTLSIMAFYLYGLWYI 150  
Db 121 aydvsvaqekmwesfekrlklhdgdgllmkaneyflfkalstlsimafylyglwyi 180

Qy 151 TSACLLALAWQOFGWLTHFCHQOQTKNRPNDTISLFFGNFLOGFSRDWKKDKHNTHTHA 210  
Db 181 tsacllalawqgfwlthefchqgqtknripndtislffgnlfgfsrdwkkdkhnthta 240

QY 211 ATNVIDHGDIDLAFAPFIPGDLCKYKASFKAALKIVPYOHLFTAMPLMLRFSWTGQ 270  
Db 241 atnvldhgdldlplafafipgdlckykasfekalkivpyghlyftamplmlrfswtgq 300  
QY 271 SVQWFKENOMEYKYVORNAFWEQATIVCHWAWVFOLFLLPTWPLRVAYFIISOMGGCL 330  
Db 301 svqwfkemqeykyvqrnatweqatlvghwawvfyqfilptwplrvaayfiisqmggsl 360  
QY 331 LIAHVVTFNHNSVDKYPANSRILNFAALQILITRNMTSPSPIDWLGGLNYQIEHHLP 390  
Db 361 liahvtfnhnsvdkypansrillnfaalqilitrnmtspfidwlgglngyqiehhlp 420  
QY 391 TPNRNLNACVYVKEWKENLPLVDDYFDGYAMNLOQLKMAEHQAQAA 443  
Db 421 tprcnlnacmkvkwckennlplylvdgyamnlqqlkmaehiqakaa 473

RESULT 3  
AAAY21891  
ID AAAY21891 standard; Protein; 447 AA.  
XX  
AC AAAY21891;  
XX  
DT 24-SEP-1999 (first entry)  
XX  
DE C. elegans delta 5-fatty acid desaturase.  
XX  
KW Delta 5-fatty acid desaturase; dihomogamma linolenic acid; gene therapy;  
KW arachidonic acid; cholesterol; polyunsaturated fatty acid; foodstuff;  
KW dietary supplement; prostaglandin.  
XX  
OS Caenorhabditis elegans.  
XX  
PN WO9933958-A2.  
XX  
PD 08-JUL-1999.  
XX  
PF 23-DEC-1998; 98WO-GB03895.  
XX  
PR 29-JUN-1998; 98GB-0014034.  
PR 23-DEC-1997; 97GB-0027256.  
XX  
PA (UYBR-) UNIV BRISTOL.  
XX  
PI Michaelson L, Napier JA, Stobart K;  
XX WPI; 1999-444067/37.  
DR N-PSDB; AAX86961.  
XX  
PT New isolated delta5-fatty acid desaturase enzymes useful in gene  
therapy  
XX  
PS Claim 18; Page 23; 36pp; English.  
XX  
CC The invention provides delta 5-fatty acid desaturases obtained from  
CC Mortierella alpina and Caenorhabditis elegans. The Delta 5-fatty acid  
CC desaturases catalyze the production of polyunsaturated fatty acids, e.g.  
CC the conversion of dihomogamma linolenic acid to arachidonic acid. The  
CC genes can be used in gene therapy as a preventative treatment, e.g. in  
CC patients suffering from high levels of cholesterol or other conditions  
CC where administration of polyunsaturated fatty acids may have beneficial  
CC disease-preventative effects. The polyunsaturated fatty acids can be used  
CC in foodstuffs or dietary supplements. The Delta 5-fatty acid desaturases  
CC can also be used for the synthesis of prostaglandins or modulation of the  
CC synthesis. The products can also be used for detection and diagnosis. The  
CC present sequence represents the C. elegans delta 5-fatty acid desaturase.  
XX  
SQ Sequence 447 AA;

Query Match 45.3%; Score 1092.5; DB 20; Length 447;  
Best Local Similarity 46.3%; Fred. No. 3.2e-99;  
Matches 201; Conservative 84; Mismatches 140; Indels 9; Gaps 4;

QY 12 MKVDGKWLYSBELVKKHPGGAVIEQYRNSDATHIFHAFHEGSSQAYKOLDLLKKHGEHD 71  
Db 13 ikidgkcwclddavlrshpggsaittykumdatvfthtgtskeayqwtelkkeepcq 72  
QY 72 E----FLEKQLERKLDKVDINVSAYDVSVAQEKWKVESEKLRQKLHDDGLMKANETFL 127  
Db 73 epeipdikdpikgid--dvnmgtfniseksaqinkstfdirmrvaeglmndgsplfyi 130  
QY 128 FKAISTLSIMAFAYLQYLGWYITTSACLLALAWOQFGLTHERFCHQOPTKNRPLNDTISL 187  
Db 131 rkiletifilfalyqyhtyylpsailngvawqqlgwlihefahqqlfknryyndlasy 190  
QY 188 FFGNFIQGSRDWKKDNHTHAATNVIDHGDIDIDLAPLFAFIPGDLCKYKASFKAALK 247  
Db 191 fvgnfiqgfsqgwkqehnhvhaatnvvgrdgdldlvfyatvaehlhny--sqdswvmt 248  
QY 248 IVPYQHLIETAMPLMLRFSWTGQSVQWVFKENOMEYKYVORNAFWEQATIVGHWAWVFYQ 307  
Db 249 ifrwqhwhvfmfpflrlslqslifvsqmpthyydyrntaiyeqvglsihwawslgq 308  
QY 308 LFLPTWPLRVAYFIISOMGGGLLIAHVVTFNHNSVDKYPANSRILNFAALQILTRNM 367  
Db 309 lyfldwstrimflvshlvvggflshvvtfnhysvekfalssninsnyaciqimtrrm 368  
QY 368 TPSPFIDWLGGLNYQIEHHLPFTMPRCNLNACVYVKEWCKENNLPLYLVDYFDGYAMN 427  
Db 369 rprgrfidwlgglngyqiehhlpftrmrlntvmpivkfeaaangipymvddyftgfwle 428  
QY 428 LOOLKMAEHQAQAA 441  
Db 429 iegfirnia-nvaak 441

RESULT 4  
AAAY96721  
ID AAAY96721 standard; Protein; 447 AA.  
XX  
AC AAAY96721;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE C. elegans fatty acid delta-5-desaturase.  
XX  
KW Fatty acid delta-5 desaturase; polyunsaturated; oil; seed;  
KW infant formula; dietary supplement.  
XX  
OS Caenorhabditis elegans.  
XX  
PN WO200034439-A1.  
XX  
PD 15-JUN-2000.  
XX  
PF 06-DEC-1999; 99WO-US28655.  
PR 07-DEC-1998; 98US-0111301.  
XX  
PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX  
PI Browse JA, Wallis JG, Watts JL;  
XX WPI; 2000-431293/37.  
DR N-PSDB; AAA51232.  
XX  
PT Purified protein having desaturase activity, useful for creating a  
PT double-bond between two carbons  
XX  
PS Claim 3; Fig 6A; 74pp; English.  
XX  
CC This is the Caenorhabditis elegans fatty acid delta-5 desaturase. The  
CC cDNA sequence is useful for recombinant production of the enzyme and for  
CC generating transformed host cells and transgenic plants. The desaturase  
CC can be used for creating a double-bond between two carbons. In

CC particular, the enzyme is useful for producing polyunsaturated fatty  
CC acids and for generating enzyme-specific antibodies useful for  
CC identifying desaturases. Oil-seed plants may be engineered to incorporate  
CC the enzyme, so that the plants produce seed oil rich in fatty acids. The  
CC fatty acids could be incorporated usefully into infant formula, foods of  
CC all kinds, dietary supplements, nutraceutical and pharmaceutical  
CC formulations.

XX Sequence 447 AA;

Query Match 45.3%; Score 1092.5; DB 21; Length 447;  
Best Local Similarity 46.3%; Pred. No. 3.2e-99;  
Matches 201; Conservative 84; Mismatches 140; Indels 9; Gaps 4;

QY 12 MKVDCGKWLVLSEELVKKHPGGAVIEQYRNSDATHIFAFHEGSSQAYKQLDLLKKHGEHD 71  
Db 13 ikidgkwcqiddavlrshpggsaittykmdattvfhftgskayqwltekkceptq 72  
QY 72 E----FLEKQLEKRLDKVDINVSAYDVSVAQEKKMVESFEKLRQKLHDDGLMKANETVFL 127  
Db 73 epeipdkddpikgid--dvnmgtfnisekrsaginkstfdlrmvrvaeglmdgsplyfi 130  
QY 128 FKAITSLIMAFAYLYQVIGWYITSAACLLALAWQFGWLTHFCHQOPTKRNPLNDTISL 187  
Db 131 rkiletifilfafylyqhtyylpsaimgvawqqlglwlihefahhqlfknyyndlasy 190  
QY 188 FPGNFGFSRDMWKDKHNTTHAATNVIDHDGIDIDLAPLFAFIPGDLCKYKASFEKAILK 247  
Db 191 fvgnflqgfsagqeqhnhvhaatcnvgrgdgldlvpfyatvaeahlnny--sqdswmwt 248  
QY 248 IVPYQHLFTAMLPMLRFSWTGSQVOWFKEKQMEYKYQYRNFAWEQATIVGHNAWVFYQ 307  
Db 249 lfwrqhvhwtfmlpflrlswllqslfvsqmpthydyrntaiyevqvglslhawsq 308  
QY 308 LFLPTPLRVAYFIISQMGGLLTAHVVTNNHNSVDKYPANSRILNNFAALQILTTNNM 367  
Db 309 lyflpdwstrimffilvshlvvgfllshvvtfnhsvekfalsnmsnyacqlqimtttrnm 368  
QY 368 FSPPIDLWGLNYQIEHHLPPTMPCRNACVYKVKEMCKENNLPLYLVDYDFGYAMN 427  
Db 369 tpgfrfidlw99lnyqiehhlfptmtrhnlntvmlpvkfeaaanglpymvddyftgfwie 428  
QY 428 LQOLKNAEHIQAK 441  
Db 429 lqefrnia-nvaak 441

RESULT 5

AA51355  
ID AAY51355 standard; Protein; 454 AA.

XX AC AAY51355;

XX 27-APR-2000 (first entry)

DE Protein 25bce with delta6 fatty acid desaturase activity.

KW sphingolipid desaturase; sld1; sphingobase; ceramide; capnoid;  
KW transgenic plant; crop plant; delta-8-unsaturated long-chain base;  
KW tolerance; resistance; soil salinity; ion stress; toxicity; drought;  
KW cold; frost; phytopathogenic microorganism; flowering time; cosmetic;  
KW pharmaceutical; food; chemical raw material.

OS Unidentified.

XX DE19828850-A1.

XX 30-DEC-1999.

XX 27-JUN-1998; 98DE-1028850.

XX 27-JUN-1998; 98DE-1028850.

XX (GVSE-) GVS GES ERWERB & VERW LANDWIRTSCHAFTLICH.

XX Heinz E, Zaehringer U, Schmidt H, Sperling P;

XX WPI; 2000-127549/12.

XX New sphingolipid desaturase that selectively introduces double bond  
XX into sphingolipids and capnoids -

XX Disclosure: Fig 16; 62pp; German.

XX This invention describes a novel sphingolipid desaturase that selectively  
XX introduces a double bond into the sphingobase of the ceramide residue of  
XX sphingolipids and capnoids. A DNA sequence encoding the sphingolipid  
XX desaturase, or a vector containing the DNA sequence, can be used to  
XX produce transgenic plants, especially crop plants, with an increased or  
XX decreased delta-8-unsaturated long-chain base content or an altered  
XX delta-8-unsaturated long-chain base cis/trans ratio, especially to  
XX compensate for a delta-8-unsaturated long-chain base deficiency, to  
XX exclude production of delta-8-unsaturated bases, to increase tolerance  
XX or resistance to soil salinity, ion stress or toxicity, drought, wet  
XX conditions, cold or frost and/or phytopathogenic microorganisms, or to  
XX alter size growth and flowering time. Cells, transgenic organisms or  
XX plants containing the DNA sequence can be used to produce sphingolipids  
XX and capnoids with unsaturated sphingobases. The sphingolipids or capnoids  
XX can be used in cosmetics, pharmaceuticals and foods and as chemical raw  
XX materials. This sequence represents a protein which has delta6 fatty acid  
XX desaturase activity which is described in the method of the invention.

XX Sequence 454 AA;

Query Match 44.8%; Score 1079; DB 21; Length 454;  
Best Local Similarity 45.6%; Pred. No. 7e-98;

Matches 201; Conservative 84; Mismatches 140; Indels 16; Gaps 5;

QY 12 MKVDCGKWLVLSEELVKKHPGGAVIEQYRNSDATHIFAFHEGSSQAYKQLDLLKKHGEHD 71  
Db 13 ikidgkwcqiddavlrshpggsaittykmdattvfhftgskayqwltekkceptq 72

QY 72 E----FLEKQLEKRLDKVDINVSAYDVSVAQEKKMVESFEKLRQKLHDDGLMKANETVFL 127  
Db 73 epeipdkddpikgid--dvnmgtfnisekrsaginkstfdlrmvrvaeglmdgsplyfi 130

QY 128 FKAITSLIMAFAYLYQVIGWYITSAACLLALAWQFGWLTHFCHQOPTKRNPLNDTISL 187  
Db 131 rkiletifilfafylyqhtyylpsaimgvawqqlglwlihefahhqlfknyyndlasy 190

QY 188 FPGNFGFSRDMWKDKHNTTHAATNVIDHDGIDIDLAPLFAFIPGDLCKYKAS 240  
Db 191 fvgnflqgfsagqeqhnhvhaatcnvgrgdgldlvpfyatvaeahlnny--s 248

QY 241 FEKAILKIVPYQHLFTAMLPMLRFSWTGSQVOWFKEKQMEYKYQYRNFAWEQATIVGH 300  
Db 249 qdswmwtlfrwqhvhwtfmlpflrlswllqslfvsqmpthydyrntaiyevqvglslh 308

QY 301 WAWVFYQLFLPTPLRVAYFIISQMGGLLTAHVVTNNHNSVDKYPANSRILNNFAALQ 360  
Db 309 wawslgqlfildwstrimffilvshlvvgfllshvvtfnhsvekfalsnmsnyacqlq 368

QY 361 ILTTRNTPSPFIDWLMGGLNYQIEHHLPPTMPCRNACVYKVKEMCKENNLPLYLVDY 420  
Db 369 imttrnmrpg-fidwlgglonyqiehhlfptmtrhnlntvmlpvkfeaaanglpymvddy 428

QY 421 FDGYAMNLOQLKNAEHIQAK 441  
Db 429 ftgfwleiegrnia-nvaak 448

RESULT 6

AA96722  
ID AAY96722 standard; Protein; 422 AA.









XX	Sequence	483 AA;	
XX	Query Match	23.28; Score 559; DB 22; Length 483;	
XX	Best Local Similarity	32.6%; Pred. No. 2e-46;	
XX	Matches	136; Conservative 66; Mismatches 155; Indels 60; Gaps 13;	
QY	18 WLYLSEEL-----VKHPCGAVTEQYRNSDATHIFHAFHEGSSQAYKQLDLLKKHGEHD	71	
DB	78 wmlvkekvydlrsfaddhpggtvistyfgrdgtvfatfppa--awkql-----nd	127	
QY	72 EFL- EKOLEKRLRDVDINVSAYDVSVAQEKKWSESFEKLRKLDHDDGLMKANETVFLFK-	129	
DB	128 yyigdlareepld-----ellkdyrdmraefvreglfksskawflq	170	
QY	130 ----AISTLSIMAFYQLYLGWIT-SACLLALAWQOFGWLTHEFCHQOQPTKNRPLNDT	184	
DB	171 llnaalfaasiatidycksy--waivlsasimgifvgcgwlahdfllhqvfenrtansf	228	
QY	185 ISLPEFGFLOGFSRDWKKDHNTHAATNVIDH-----DGDIDLAPLFAFPGDLCKYKA	239	
DB	229 fgyllfgncvlgfsvswwrtknhlhtspnecdegytldedidtlplawskellatves	288	
QY	240 SFERAILKIVPYQHLYFTAMLPMLRFSGVSWGVFKENOMEXKVYQRNAFWQATIVG	299	
DB	289 --krilrvlqyqhymlpllfmaryswtfgsliftfnpdlsttk----gliekgtvaf	340	
QY	300 HWAWVFQQLP-LLPTWPLRVAYFIISOGGGLLTAHVVTFNHNSVDKYPANSRILNNFAA	358	
DB	341 hyawfswaathilpgvakplawmvatelvaglllgfvtishngkevynes----kdfvr	396	
QY	359 LQILFTNRMTSPSPIDMLWGLNLYQIEHHLPFTMPRCNLNACVYKVECKENNLPLY	415	
DB	397 aqvlttrntrkgwfnwftggldtqiehhiftprhnykpiapqvealckkhgley	453	
XX	RESULT 11		
XX	AAB46436		
XX	ID	NAB46436 standard; Protein; 483 AA.	
XX	AAB46436;		
XX	06-APR-2001 (first entry)		
XX	DT		
XX	DE	C. purpureus delta6-acetylenase/delta6-desaturase SEQ ID NO 4.	
XX	KW	Delta6-acetylenase; delta6-desaturase; unsaturated fatty acid;	
XX	KW	transgenic plant; plant oil; triglyceride; nutrition; animal feed;	
XX	KW	cosmetic.	
XX	OS	Ceratodon purpureus.	
XX	OS	W0200075341-A1.	
XX	PN	14-DEC-2000.	
XX	PD		
XX	PF	07-JUN-2000; 2000WO-EP05274.	
XX	PR	07-JUN-1999; 99DE-1025718.	
XX	PR	22-DEC-1999; 99DE-1062409.	
XX	XX		
XX	XX	(BADI ) BASF AG.	
XX	PI	Heinz E, Stymne S, Lee M, Girke T, Sperling P, Zaehringer U;	
XX	PI	WPI; 2001-112150/12.	
XX	DR	N-PSDB; AAF25730.	
XX	DR		
XX	XX		
XX	PT	Nucleic acid encoding delta6-acetylenase or desaturase, useful for	
XX	PT	producing plant oils with increased content of unsaturated fatty acids	
XX	PT	-	
XX	XX		

```

XX OS Mortierella alpina.
XX PN WO9846763-A1.
XX PD 22-OCT-1998.
XX PF 10-APR-1998; 98WO-US07126.
XX PR 11-APR-1997; 97US-0834655.
XX PA (ABBO ) ABBOTT LAB.
XX FA (CALJ ) CALGENE LLC.
XX PI Chaudhary S, Huang Y, Knutzon D, Leonard AE, Mukerji P;
XX PI Thurmond J;
XX DR WPI; 1998-594582/50.
XX DR N-PSDB; AAV63624.
XX PT New isolated fatty acid desaturase enzymes - used for the production
XX PT of polyunsaturated fatty acids for use in, e.g. pharmaceutical
XX PS compositions, nutritional compositions, cosmetics or animal feed
XX PS Claim 3; Fig 3A-E; 165pp; English.
XX CC The present sequence represents a Mortierella alpina fatty acid delta-6
XX CC desaturase enzyme. The enzyme sequence is used in the methods of
XX CC the invention. The specification describes methods for desaturating a
XX CC fatty acid and for producing a desaturated fatty acid by expressing
XX CC increased levels of a desaturase. The present desaturase is an enzyme
XX CC which introduces a double bond carbons 6 and 7 from the carboxyl end of
XX CC a fatty acid molecule. The enzyme can be used for desaturating fatty
XX CC acids. The enzyme can be used to produce polyunsaturated fatty acids,
XX CC which can be used for treating malnutrition, in pharmaceutical
XX CC compositions, in cosmetics or in animal feed. The polyunsaturated fatty
XX CC acids can be used for treating e.g. restenosis after angioplasty,
XX CC inflammation, rheumatoid arthritis, asthma, psoriasis, cancer, diabetes
XX CC or eczema or reduce blood pressure. They can also be used to inhibit
XX CC platelet aggregation, cause vasodilation, lower cholesterol levels,
XX CC inhibit proliferation of vessel wall smooth muscle and fibrous tissue,
XX CC reduce or prevent gastro-intestinal bleeding and other side effects
XX CC caused by non-steroidal anti-inflammatory drugs, prevent or treat
XX CC endometriosis and premenstrual syndrome, treat myalgic encephalomyelitis
XX CC and chronic fatigue after viral infections, treat AIDS, multiple
XX CC sclerosis, acute respiratory syndrome, hypertension and inflammatory skin
XX CC disorders.
XX SQ Sequence 457 AA;

Query Match 23.1%; Score 557; DB 19; Length 457;
Best Local Similarity 32.1%; Pred. No. 2.9e-46;
Matches 143; Conservative 64; Mismatches 187; Indels 52; Gaps 11;

Qy 12 MKVCGKWLSEELVKHGGAVIEQYRNSDATHIFAHFEGSSOAYKQLDLLKKGHSD 71
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 34 mldnk-vydvrefvdpdpggsvllthvgkdgtdvftfhpeaaw----- 77

Qy 72 EFLEKOLEKRLDKVDINVSAYDVSAQEKWVESFEKLRQKLHDDGLMKANETFLFKAI 131
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 78 etlanfyvgdidesdrdknddfa-aevrkirtlfqsl-----gyddskayafkvs 129

Qy 132 STLSTMAFAFYLOVLGW-----YITSACLLALAWQOFGWLTHFCHQOQTPKRNPLNDT 184
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 130 fnlciglstvi-vakvgqgtstlanvlsaaallgfwcqcgwlandflhhqvfqdrfwgd 188

Qy 185 ISLFGFNPLQGSRDWKKDKHNTHTAATNVIDHGDIDIDLAPL-----FAFIG-D 233
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 189 fgafgvcgqgfswwkdkhnhhaapnvhgedpdidhplltwsehalcmfsdvpdee 248

Qy 234 LCKYKASFEKAILKIVPYQHLYFTAMLPMLRFSWTGQSVQVWFENQMEYKVYQR--NAF 291
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 249 ltrmwsrf-----mvlngtwyfipillsfarlswclqslflvlpngqahkpsgarvpisl 302
Qy 292 WEQATIVGHWAWFYQLFLPTWPLR-VAYFIISQMGGLLIAHVVTFNHNSVDKYPANS 350
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 303 veqlslamhwtwlatmflfkdpvnmivfyflvsqavcqnllaivfslnhngmpviskee 362
Qy 351 RILNFAALQILTTTRMTPTSPFIDWLWGLNVOIEHHLPFTMPRNLNACVKYKWECKE 410
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 363 avdmfftkqitgrdvhpglfanwftgglnyqiehlflpsmrhnsfkqpavetlokk 422
Qy 411 NNPLPYLVDDYFDGYAMNLOOLKNMAE 436
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 423 ynvryhttgmiectaevfsrlnevs 448

RESULT 13
AAW95504
ID AAW95504 standard; peptide; 457 AA.
XX AC AAW95504;
XX DT 26-MAR-1999 (first entry)
XX DE Mortierella alpina delta 6 desaturase.
XX KW Delta 6 desaturase; recombinant; fatty acid desaturase; FAD; PUFA; oil;
XX KW polyunsaturated fatty acid; linoleic acid; arachidonic acid; linolenic;
XX KW stearidonic acid; eicosapentaenoic acid; malnutrition; feeding formula;
XX KW dietary supplement; prostaglandin; restenosis; angioplasty; inflammation;
XX KW rheumatoid arthritis; psoriasis; osteoporosis; cancer; eczema; AIDS;
XX KW diabetes; cosmetic; animal feed.
XX OS Mortierella alpina.
XX PN WO9846764-A1.
XX PD 22-OCT-1998.
XX PF 10-APR-1998; 98WO-US07421.
XX PR 24-OCT-1997; 97US-0956985.
XX PR 11-APR-1997; 97US-0833610.
XX PR 11-APR-1997; 97US-0834033.
XX PR 11-APR-1997; 97US-0834655.
XX PA (ABBO ) ABBOTT LAB.
XX PA (CALJ ) CALGENE LLC.
XX PI Chaudhary S, Huang Y, Knutzon D, Leonard AE, Mukerji P;
XX PI Thurmond J;
XX DR WPI; 1999-080739/07.
XX DR N-PSDB; AAX00889.
XX PT Nucleic acid construct able to express fatty acid desaturase in
XX PT plants - useful in human or animal nutrition, as cosmetics and
XX PT therapeutically, e.g. for restenosis, cancer and diabetes
XX PS Claim 7; Fig 3A-E; 210pp; English.
XX CC This represents a Mortierella alpina delta 6 desaturase. The invention
XX CC relates to a nucleic acid construct that contains at least one of the
XX CC nucleotide sequences (AAX00889 to AAX00891) encoding M. alpina delta 6,
XX CC delta 12 and delta 5 desaturases (AAW95504 to AAW95506) respectively,
XX CC coupled to an expression control sequence functional in plants.
XX CC Recombinant plant cells containing at least one DNA encoding a M. alpina
XX CC fatty acid desaturase (FAD), can be used for the production of
XX CC polyunsaturated fatty acid (PUFA). These recombinant cells or plants
XX CC containing them are used to produce oils such as linoleic acid,
XX CC arachidonic acid, gamma-linolenic acid, dihomo-gamma-linolenic acid,
XX CC stearidonic acid and eicosapentaenoic acid (EPA). These plant oils are
XX CC used: (i) to treat malnutrition; (ii) in infant feeding formulas, or
XX CC dietary supplements or substitutes, for use in humans or animals; (iii)

```

CC for treating disorders associated with inadequate consumption or  
CC production of PUFA (or their metabolites such as prostaglandins), e.g.  
CC restenosis after angioplasty, inflammation, AIDS, rheumatoid arthritis, and  
CC psoriasis, osteoporosis, cancer, eczema, diabetes; (iv) as cosmetics, and  
CC (v) as animal feeds. Fragments of the DNA are used as probes to isolate  
CC related coding sequences. Recombinant plants can produce high yields of  
CC PUFA, since new pathways can be created and unwanted ones suppressed.  
CC Plants can be engineered to express oils of particular PUFA composition,  
CC e.g. one similar to that in human milk, and product recovery is simpler  
CC than with e.g. fish.

XX  
SQ Sequence 457 AA;

Query Match 23.1%; Score 557; DB 20; Length 457;  
Best Local Similarity 32.1%; Pred. No. 2.9e-46;  
Matches 143; Conservative 64; Mismatches 187; Indels 52; Gaps 11;

Qy 12 MKVDGKWLSEELVKKHGGAVIEQYRNSDATHIFAFHEGSSQAYKQLDLLKKHGEHD 71  
Db 34 mlidnk-vydvrefvdpdpgsvilthvgkdgtdvdfthpeaaw-----gyydsakayafkvs 77  
Qy 72 EFLEKLEKRLDKVDINVSAYDSVAQEKMKVESFEKLRQKLDHDDGLMKANETFLFKAI 131  
Db 78 etlanfyvgdidesdrdknddfa-aevrkrltlfqsl-----gyydsakayafkvs 129  
Qy 132 STLSIMAFAYLQYLGW-----YITSACLLALAWQFGWLTHFCHQOQPTKRNPLNDT 184  
Db 130 fnlcwglstvi-vakwgqstlanvisaalglfwqcgwlahdflhhqvqgrfwdl 188  
Qy 195 ISLFFGNFLOGFSRDWKKKHNTHHAATNVTDHGDIDLAPL-----FAFIPG-D 233  
Db 189 fgafllgvcgqfsswwkdkhthhaapnvhgedpdidthplltwsehalcmfsvdpee 248  
Qy 234 LCKYKASFEKAILKIVPYQHLYFTAMPLMPLRFSWTGQSVQVFKENQMEYKVYQOR--NAF 291  
Db 249 ltrmwsrf-----mvlntwfyfpilsfarslswclsgsilvlpngahkpsgarvpsl 302  
Qy 292 WEQATIVGHAWVFYQLFLPTWPLR-VAYFTISQMGGLLIAHVTFNHNVSVDKYPANS 350  
Db 303 veqlslamhwtlatmflikdpvnmlyflvsqvcgnllaivfslnhngmpviskee 362  
Qy 351 RILNFAALQILTRNMTSPFIDWLNGGLNYQIEHHLFPMPRCNLNACVKKYKWECKE 410  
Db 363 avdmdfckqitgrdvhpgifanfwfvgglnyqiehlfpmsprhntskiqpavetlckk 422  
Qy 411 NNLPLYLVDYDYGAMNQLQKNMAE 436  
Db 423 ynvryhtgmiegtaevfsvrlnvsk 448

RESULT 14  
AA92599  
ID AA92599 standard; Protein; 457 AA.

AC  
XX

AC  
XX

DT 10-AUG-2000 (first entry)

DE M. alpina delta-6 fatty acid desaturase.

XX delta-6 desaturase; gamma-linolenic acid; biosynthesis;

KW transgenic insect cell; polyunsaturated long chain fatty acid;

KW antiinflammatory; antirheumatic; antiasthmatic; antipsoriatic;

KW osteopathic; cytostatic; anti-diabetic; dermatological; gynecological;

KW anti-HIV; neuroprotective; hypotensive; nephrotropic; vasodilator;

KW antiaggregant; vasotropic.

XX Mortierella alpina.

OS

XX

FX Key Location/Qualifiers

FT Misc-difference 172..176

FT /label= histidine\_box

FT Misc-difference 209..213  
FT /label= histidine\_box  
FT /note= "conserved among membrane-bound desaturases"  
FT Misc-difference 395..399  
FT /label= histidine\_box  
FT /note= "conserved among membrane-bound desaturases"  
XX WO200020602-A2.

XX 13-APR-2000.

XX 29-SEP-1999; 99WO-US22686.

XX 05-OCT-1998; 98US-0103110.

XX (ABBO ) ABBOTT LAB.

XX Mukerji P, Huang Y, Parker-barnes JM, Das T;

XX WPI; 2000-328935/28.

XX N-PSDB; AAA09430.

XX Novel transgenic insect cells comprising a nucleotide sequences which

XX encode delta-6-desaturase or delta-12- desaturase, useful for producing

XX poly-unsaturated long chain fatty acids, e.g. arachidonic acid

XX Claim 1; Page 145-146; 170pp; English.

XX The fatty acid desaturases are able to catalyse the conversion of oleic

XX acid to linoleic acid, linoleic acid to gamma-linolenic acid or of

XX alpha-linolenic acid to stearidonic acid. Transgenic insect cells

XX comprising a nucleotide sequence which encodes a polypeptide comprising

XX residues 50-53, 39-43, 172-176, 204-213, or 390-402 of delta-6

XX desaturase (AA92599) or comprising delta-12 desaturase (AA92600) are

XX claimed. Oil and fatty acids (especially gamma-linolenic acid) isolated

XX from the recombinant insect cells are also claimed. Production of

XX polyunsaturated long chain fatty acids (PUFAs) in insect cells has many

XX advantages, as insect cells have greatly simplified lipid compositions,

XX are not subject to external variable fluctuations, and can easily be

XX maintained and manipulated. The oils are used in pharmaceutical

XX compositions, infant formulas, dietary supplements or substitutes, and

XX cosmetics (all claimed). The PUFA supplements have antiinflammatory,

XX antirheumatic, antiasthmatic, antipsoriatic, osteopathic, cytostatic,

XX antidiabetic, dermatological, gynecological, anti-HIV, neuroprotective,

XX hypotensive, nephrotropic, vasodilator, antiaggregant and vasotropic

XX activity.

XX SQ Sequence 457 AA;

Query Match 23.1%; Score 557; DB 21; Length 457;

Best Local Similarity 32.1%; Pred. No. 2.9e-46;

Matches 143; Conservative 64; Mismatches 187; Indels 52; Gaps 11;

Qy 12 MKVDGKWLSEELVKKHGGAVIEQYRNSDATHIFAFHEGSSQAYKQLDLLKKHGEHD 71

Db 34 mlidnk-vydvrefvdpdpgsvilthvgkdgtdvdfthpeaaw-----gyydsakayafkvs 77

Qy 72 EFLEKLEKRLDKVDINVSAYDSVAQEKMKVESFEKLRQKLDHDDGLMKANETFLFKAI 131

Db 78 etlanfyvgdidesdrdknddfa-aevrkrltlfqsl-----gyydsakayafkvs 129

Qy 132 STLSIMAFAYLQYLGW-----YITSACLLALAWQFGWLTHFCHQOQPTKRNPLNDT 184

Db 130 fnlcwglstvi-vakwgqstlanvisaalglfwqcgwlahdflhhqvqgrfwdl 188

Qy 185 ISLFFGNFLOGFSRDWKKKHNTHHAATNVTDHGDIDLAPL-----FAFIPG-D 233

Db 189 fgafllgvcgqfsswwkdkhthhaapnvhgedpdidthplltwsehalcmfsvdpee 248

Qy 234 LCKYKASFEKAILKIVPYQHLYFTAMPLMPLRFSWTGQSVQVFKENQMEYKVYQOR--NAF 291

Db 249 ltrmwsrf-----mvlngtqfwfypilslfarslswclqsilflvlpngqahkpsgarvpisl 302  
QY 292 WEQATIVGHAWVYQQLFLLPTWPLR-VAYFTIISQMGGLLIAHVVTNHNHNSVDKYPANS 350  
Db 303 veqlslamhwtwylatmflfikdpvnmlyflvsqavcgnliaivfslnhngmpviskee 362  
QY 351 RILNFAALQILTRNMTSPFIDMLWGLNYQIEHHLFPTMPRCNLNACVKYKEMCKE 410  
Db 363 avdmddfktqitgrdvhpglfanwftgglngyqiehhlfpsmprhnfskiqavetlckk 422  
QY 411 NNLPLYVDDYFDGYAMNLOQLKNAE 436  
Db 423 ynvryhtgmiegtaevfslnevs 448

## RESULT 15

AAV56045  
ID AAV56045 standard; Protein; 457 AA.

XX AAV56045;

XX 28-MAR-2000 (first entry)

XX Fungal delta6-desaturase protein sequence.

XX Polyunsaturated fatty acid; fungus; delta6-desaturase; animal feed;  
KW transgenic animal; malnutrition; biosynthesis.

XX Unidentified.

XX WO9961602-A1.

XX 02-DEC-1999.

XX 28-MAY-1999; 99WO-US12088.

XX 29-MAY-1998; 98US-0087578.

XX (OHIS ) UNIV OHIO STATE.

XX Kopchick JJ, Kelder B, Huang Y, Kirchner SJ, Mukerji P;

XX WPI; 2000-072619/06.

DR N-PSDB; AAZ47129.

XX Producing essential fatty acids and long-chain polyunsaturated fatty  
PT acids, for use in nutritional, animal feed and medical formulations -

XX Disclosure; Fig 9; 71pp; English.

XX The invention relates to a method of generating novel compositions  
CC comprising animal cells producing essential fatty acids (FAs). The animal  
CC cells are produced by transforming cells, e.g. embryonic stem cells, with  
CC nucleic acid encoding heterologous enzymes involved in fatty acid,  
CC e.g. long chain or polyunsaturated fatty acid (PFA) biosynthesis. This  
CC sequence corresponds to a fungal delta6-desaturase whose coding sequence  
CC is an example of a nucleic acid sequence used to transform the cells. The  
CC essential FAs obtained can be used in nutritional formulations or animal  
CC feed formulations. The long chain PFAs can be used in nutritional  
CC formulations, cosmetic formulations or animal feed formulations. The  
CC products can also be used for producing transgenic animals which can be  
CC used for producing essential FAs which can be used for producing  
CC downstream products such as leukotrienes, thromboxanes, arachidonic acid,  
CC eicosapentaenoic acid or docosahexaenoic acid. The products can also be  
CC used in cell culture. The animal or milk fat produced can be administered  
CC to treat malnutrition.

XX Sequence 457 AA;

Query Match 23.1%; Score 557; DB 21; Length 457;  
Best Local Similarity 32.1%; Pred. No. 2.9e-46;  
Matches 143; Conservative 64; Mismatches 187; Indels 52; Gaps 11;

QY 12 MKVDGKWLVLSEELVKHPGGAVIEQYRNSDATHIFHAFHEGSSQAYKQLDLLKKHGEHD 71  
Db 34 midnk-vydvrefpdpbggsvilthvgkdgtdvdfthpeaaw----- 77  
QY 72 EFLEKOLEKRLDKVDINVSAYDVSVAQEKMMVESFEKLRQKLHDDGLMKANETYFLFKAI 131  
Db 78 etlanfyvgdidesdrdiknddfa-aevrklrtlfqsl-----gyydsakayafkvs 129  
QY 132 STLSIMAFAYLOYLGM-----YITSACLLALAWQOFGWLTHERFCHOOPTKNSPLNDT 184  
Db 130 fnlcwglstvi-vakwqgtstlanvlsaallglfwqcgwiahdfllhqvfqdrfwgdl 188  
QY 185 ISLFFGNFLQGFSDWWDKDNTHHAATNVIDHDGIDIDLAPL-----FAFIPG-D 233  
Db 189 fgafllggvcqgfsfsswwkdkhnhhaapnvhgdpdidthplltwsehalenmfsvpdee 248  
QY 234 LCKYKASFEKAILKIVPYOHLVFTAMLPMRLFSWTGQSVQWVFKENQMEYKVYQR--NAF 291  
Db 249 ltrmwsrf-----mvlngtqfwfypilslfarslswclqsilflvlpngqahkpsgarvpisl 302  
QY 292 WEQATIVGHAWVYQQLFLLPTWPLR-VAYFTIISQMGGLLIAHVVTNHNHNSVDKYPANS 350  
Db 303 veqlslamhwtwylatmflfikdpvnmlyflvsqavcgnliaivfslnhngmpviskee 362  
QY 351 RILNFAALQILTRNMTSPFIDMLWGLNYQIEHHLFPTMPRCNLNACVKYKEMCKE 410  
Db 363 avdmddfktqitgrdvhpglfanwftgglngyqiehhlfpsmprhnfskiqavetlckk 422  
QY 411 NNLPLYVDDYFDGYAMNLOQLKNAE 436  
Db 423 ynvryhtgmiegtaevfslnevs 448

Search completed: August 6, 2002, 16:21:54  
Job time: 1063 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 6, 2002, 16:16:26 ; Search time 15.83 Seconds  
(without alignments)  
683.546 Million cell updates/sec

Title: US-09-555-093-2

Perfect score: 2410

Sequence: 1 MYVDKNASGLRMKVDGKWL.....YAMNLOQLKNMAEHIOAKAA 443

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	557	23.1	457	2	US-08-834-655-2
2	557	23.1	457	3	US-08-834-033A-2
3	557	23.1	457	4	US-09-363-574-2
4	555	23.0	457	2	US-08-833-610-4
5	555	23.0	457	3	US-08-834-033A-14
6	501	20.8	355	2	US-08-834-655-5
7	501	20.8	355	3	US-08-834-033A-6
8	501	20.8	355	4	US-09-363-574-5
9	456.5	18.9	446	2	US-08-833-610-5
10	456.5	18.9	446	3	US-08-834-033A-15
11	448.5	18.6	448	1	US-08-366-779-5
12	448.5	18.6	448	1	US-08-789-936-5
13	448.5	18.6	448	4	US-08-934-254-5
14	412	17.1	87	2	US-08-834-655-10
15	412	17.1	87	3	US-08-834-033A-11
16	412	17.1	87	4	US-09-363-574-10
17	400	16.6	452	4	US-08-934-254-27
18	259.5	10.8	252	2	US-08-834-655-7
19	259.5	10.8	252	3	US-08-834-033A-8
20	259.5	10.8	252	4	US-09-363-574-7
21	217	9.0	365	2	US-08-833-610-7
22	217	9.0	365	3	US-08-834-033A-17
23	215.5	8.9	359	1	US-08-307-382-2
24	215.5	8.9	359	1	US-08-366-779-2
25	215.5	8.9	359	1	US-08-478-727-2
26	215.5	8.9	359	1	US-08-473-508-2
27	215.5	8.9	359	1	US-08-789-936-2

28	215.5	8.9	359	2	US-08-833-610-6	Sequence 6, Appli
29	215.5	8.9	359	3	US-08-834-033A-16	Sequence 16, Appl
30	215.5	8.9	359	4	US-08-934-254-2	Sequence 2, Appli
31	186	7.7	446	2	US-08-833-610-2	Sequence 2, Appli
32	186	7.7	446	3	US-08-834-033A-5	Sequence 5, Appli
33	174.5	7.2	443	4	US-09-161-994A-3	Sequence 3, Appli
34	166	6.9	143	2	US-08-834-655-11	Sequence 11, Appl
35	166	6.9	143	3	US-08-834-033A-12	Sequence 12, Appl
36	166	6.9	143	4	US-09-363-574-11	Sequence 11, Appl
37	164.5	6.8	418	5	PCT-US94-01321-72	Sequence 72, Appl
38	161	6.7	131	2	US-08-834-655-9	Sequence 9, Appli
39	161	6.7	131	3	US-08-834-033A-10	Sequence 10, Appl
40	161	6.7	131	4	US-09-363-574-9	Sequence 9, Appli
41	149	6.2	125	2	US-08-834-655-8	Sequence 8, Appli
42	149	6.2	125	3	US-08-834-033A-9	Sequence 9, Appli
43	149	6.2	125	4	US-09-363-574-8	Sequence 8, Appli
44	148	6.1	424	4	US-09-161-994A-4	Sequence 4, Appli
45	141.5	5.9	384	4	US-09-354-231B-18	Sequence 18, Appl

## ALIGNMENTS

RESULT 1

US-08-834-655-2

; Sequence 2, Application US/08834655

; Patent No. 5968809

; GENERAL INFORMATION:

; APPLICANT: KNUTZON, DEBORAH

; APPLICANT: MURKERJL, PRADIP

; APPLICANT: HUANG, YUNG-SHENG

; APPLICANT: THURMOND, JENNIFER

; APPLICANT: CHAUDHARY, SUNITA

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS

; TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RAB-VENTER LAW GROUP, P.C.

; STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039

; CITY: PALO ALTO

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/834.655

; FILING DATE: 11-APR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: RAE-VENTER, BARBARA

; REGISTRATION NUMBER: 32,750

; REFERENCE/DOCKET NUMBER: CGNE.124.0005

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 328-4400

; TELEFAX: (650) 328-4477

; TELEX: N/A

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 457 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-834-655-2

Query Match

Best Local Similarity 23.1%; Score 557; DB 2; Length 457;

Matches 143; Conservative 64; Mismatches 187; Indels 52; Gaps 11;

QY	12	MKVDGKWLILSBEVLVKKHPCGAVIEQYRNSDATHIFHAFHEGSSQAYKQDILLKXGHD	71
Db	34	MIIDNK-VYDREVPDHPGGSVILTHVGKDGDFDTEHPEAAW	77
QY	72	EFLEKQLEKRLKVDINVSAYDSVAQEKMYVESFEKLQKLHDDGLMKANETYFLFKAI	131
Db	78	ETLANFYVGDI DESDRIDKNDFA-AEVRKLRTLFQSL-----GYDSSKAYAFKVS	129
QY	132	STLSINAFAPYLOLW-----YITSACLALAWQCGWLTHBFCHOOPTKRNPLNDT	184
Db	130	FNLCIMGLSTVI-VAKWGOTSTLIANYLSAALLGLFWQCGWLAHDFLHHQVQDFRWGDL	188
QY	185	ISLFFGNFLQFSRDWKKHKNTHHAATNVIDHGDIDLAPL-----FAFIPG-D	233
Db	189	FGAFLGGVCQGSSSWKKDKHNTHAAPNVHGCDPDIDTHPLLTWSEHALEWFSVDPEE	248
QY	234	LCKYKASFEKAIKIVPYQHILFTAMLPMLRSWTQCSQVOWFEKENQMEYKYQR--NAN	291
Db	249	LTRWRSF-----MYLNOTWFYFPLSFARUSWCLOSILFLVPLNGQAHKPSGAKVPISL	302
QY	292	WEQATIVGHAWFYOLFLLPTWPLR-VAYFIISQGGGLLAHVYTFHNHNSVDVKYKANS	350
Db	303	VEQLSLAMHWTYLATMELFIKDPVKNLVYFLVSQAVCGNLLAIVFSLNHNGMPVISKEE	362
QY	351	RILNNEFAQILITRNWTFSPFIDMLWGLNGVIOIEHLEPTPRCNLNACVYKWECKE	410
Db	363	AVDMDFTKOIIITGRDVHGLFANWTFGTGUNTQIOEHLFPSPMRHNFPSKIQPAVETLCK	422
QY	411	NNLPYLVDYDFCYAMNLOOLKNAAE	436
Db	423	YNYRYHTGMIEGTAEVFSRLNEVSK	448

## RESULT 2

```

US-08-834-033A-2
; Sequence 2, Application US/08834033A
; Patent No. 6075183
; GENERAL INFORMATION:
; APPLICANT: KNUTZON, DEBORAH
; APPLICANT: MUKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LIMBACH AND LIMBACH, L.L.P.
; STREET: 2001 FERRY BUILDING
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,033A
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WARD, MICHAEL R.
; REGISTRATION NUMBER: 38,651
; REFERENCE/DOCKET NUMBER: CGAB-300.USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 433-4150
; TELEFAX: (415) 433-8716
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```





APPLICANT: MUKERJI, PRADIP  
APPLICANT: HUANG, YUNG-SHENG  
APPLICANT: THURMOND, JENNIFER  
APPLICANT: CHAUDHARY, SUNITA  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS  
TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LIMBACH AND LIMBACH, L.L.P.  
STREET: 2001 FERRY BUILDING  
CITY: SAN FRANCISCO  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
FILING DATE: 11-APR-1997  
APPLICATION NUMBER: US/08/834,033A  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: WARD, MICHAEL R.  
REGISTRATION NUMBER: 38,651  
REFERENCE/DOCKET NUMBER: CGAB-300.USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 433-4150  
TELEFAX: (415) 433-8716  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 457 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-834-033A-14

Query Match 23.0%; Score 555; DB 3; Length 457;  
Best Local Similarity 32.1%; Pred. No. 1.2e-46;  
Matches 143; Conservative 64; Mismatches 187; Indels 52; Gaps 11;

QY 12 MKVDGKWLILSEELVKKHGGAVIEQYRNSDATHIEHAFHEGSSQAYKQLDLLKKHGEHD 71  
DB 34 MIDNK-VYDVREVPDPHGGSVILTHVGRDGTVDFTFPEAAW----- 77  
QY 72 EFLEKQLEKRLDKYDINVSAYDVSVAQEKMKVESFEKLRQKLHDDGLMKANETYFLFKRAI 131  
DB 78 ETLANFYVGDIDESDRDIKNDDFA-AEVRKLTFLQSL-----GYDSSKAYIAFKVS 129  
QY 132 STLIMAFAYLYQLGW-----YITSACLLALAWQOFGWLTHEFCHQOQTKNRPLNDT 184  
DB 130 FNLICWGLSTVI-VAKWGQSTLANVLNSAALLGLFWQCGWLAHDFLHHQVDFRFGDL 188  
QY 185 ISLFFGNFLOGFSRDWKKHNTHAATNVIDHGDIDLAPL-----FAFIG-D 233  
DB 189 FGAFLLGGVCGQFSSWKKHNTHAAPNVEDPDIDTHPLLWSEHALEMFSDVPEDE 248  
QY 234 LCKYKASFKAAILKIVPYQHLYFTAMPLMPLRFSWTGOSVQWVFKENQMEYKVYQ--NAF 291  
DB 249 LTRWRSF-----MVLNQTWFYPILSFARLSWCLQSILFVLPNGQAHKPSGARVPISL 302  
QY 292 WEQATIVGHNAWVYQLFLPTWPLR-VAYFIISOMGGGLLIAHVTFNHNVDKYPANS 350  
DB 303 VEQLSLAMHTWYLATMFLFKIDPNNMLVFLVSOAVCGNLLAIVFLSNHNGMPVISKEE 362  
QY 351 RILANFAALQILTRNTPSPFIDMLGGLNLYQIEHLFPTMPRCNLNACVKKVKEWCKE 410  
DB 363 AVDMDFTKIITGRDVRHPLGFANWFTGGGLNYQIEHLFLPSMPRNFNSKIQPAVETLCKK 422

QY 411 NNLPYLVDYDFGYAMNLLQOLKNMAE 436  
DB 423 YNRYHTTGMTEGTAEVESRLNEYSK 448  
RESULT 6  
US-08-834-655-5  
Sequence 5, Application US/08834655  
Patent No. 5968809  
GENERAL INFORMATION:  
APPLICANT: KNUZON, DEBORAH  
APPLICANT: MUKERJI, PRADIP  
APPLICANT: HUANG, YUNG-SHENG  
APPLICANT: THURMOND, JENNIFER  
APPLICANT: CHAUDHARY, SUNITA  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS  
TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: RAE-VENTER LAW GROUP, P.C.  
STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,655  
FILING DATE: 11-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: RAE-VENTER, BARBARA  
REGISTRATION NUMBER: 32,750  
REFERENCE/DOCKET NUMBER: CGNE.124.0005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 328-4400  
TELEFAX: (650) 328-4477  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 355 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-834-655-5  
Query Match 20.8%; Score 501; DB 2; Length 355;  
Best Local Similarity 34.3%; Pred. No. 1.8e-41;  
Matches 120; Conservative 48; Mismatches 154; Indels 28; Gaps 7;  
QY 108 KLKQKLDHDDGLMKANETYFLFKALSTLSIMAFAYLYQLGW-----YITSACLLALAW 160  
DB 4 KLRTLFQSLGYDSSKAYIAFKVSFNLICWGLSTVI-VAKWGQSTLANVLNSAALLGLFW 62  
QY 161 QOFGWLTHEFCHQOQTKNRPLNDTISLFFGNFLOGFSRDWKKHNTHAATNVIDHGD 220  
DB 63 QOCCWLAHDFLHHQVDFRFGDLFGAFLLGGVCGQFSSWKKHNTHAAPNVHGEDPD 122  
QY 221 IDLAPL-----FAFIG-DLCKYKASFKAAILKIVPYQHLYFTAMPLMPLRFSWTG 269  
DB 123 IDTHPLLWSEHALEMFSDVPEDELTWMSRF-----MVLNQTWFYPILSFARLSWCL 176  
QY 270 QSVQWVFKENQMEYKVYQ--NAFWEQATIVGHNAWVYQLFLPTWPLR-VAYFIISOM 326  
DB 177 QSILFVLPNGQAHKPSGARVPISLVEQLSLAMHTWYLATMFLFKIDPNNMLVFLVSOA 236  
QY 327 GGGLLIAHVTFNHNVDKYPANSRILNFAALQILTRNTPSPFIDMLGGLNLYQIEH 386

[illegible]

```

Db 177 QSLFVLPNGQAHKSGARVPISLVEQLSLAMHTWVYLATMFLFIKDPVNMVLVFLVSOA 236
Qy 327 GGGLLIAHVTFPHNSVDKYPANSRILNNFAALQILITTRNMTSPSPFIDLWGLNQIEH 386
Db 237 VCGNLLAIVSLNHNMGMPVISKEEAVDMDFTKQIITGRDVHPGLFANFTGCLNQIEH 296
Qy 387 HLFPTMPCRNLCACVYKWKCKENNLPLYLDDYFDGCGYAMNLOQLKMAE 436
Db 297 HLFPSMPRHNSFKIOPAVETLCKKYNVRYHTTGMIETGAEVESRLNEVSK 346

RESULT 8
US-09-363-574-5
; Sequence 5, Application US/09363574
; Patent No. 6136574
; GENERAL INFORMATION:
; APPLICANT: KNUITZON, DEBORAH
; APPLICANT: MURKERJL, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LIMBACH AND LIMBACH L.L.P.
; STREET: 2001 FERRY BUILDING
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/363.574
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WARD, MICHAEL R.
; REGISTRATION NUMBER: 38,651
; REFERENCE/DOCKET NUMBER: CGAB-202 USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 433-4150
; TELEFAX: (415) 433-8716
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-363-574-5

Query Match 20.8%; Score 501; DB 4; Length 355;
Best Local Similarity 34.3%; Pred. No. 1.8e-41;
Matches 120; Conservative 48; Mismatches 154; Indels 28; Gaps 7

Qy 108 KLRKRLDDGLGMKANETFLPKAISTLSIMAFVLYQLGW-----YITSACLALAW 160
Db 4 KLRTLFQSLGYDSSKAYAFKVSFNLCLWGLSTVI-VAKWGQTSTLANVLSALLGLFW 62
Qy 161 QQFGLLTHFEHQOQTKRPLNDTISLFEGNFLOGFSRDWKKDKHNTHHAATNVIDHGD 220
Db 63 QQCGLWHLDFLHHQVQDFRWDGLFCAGFLGGVQCQGFSSSWKKDKHNTHHAAPNVHGEDPD 122
Qy 221 IDLAPL-----FAPIG-DICKYKASFEKAILKIVPQHLVFTMLPMLRFSWTG 269
Db 123 IDTHPLLWSEHALBMFSDVPDELTBMWSRF-----NVLNQTWFFPILSPARLSWCL 176

```

Query Match	20.8%	Score 501;	DB 3;	Length 355;
Best Local Similarity	34.3%;	Pred. No. 1.8e-41;		
Matches 120; Conservative	48;	Mismatches 154;	Indels 28;	Gaps 7;
QY	108	KLROKRLHDDGLMKANETVFLFKALSTLSIMAFAYFLQYLGW-----YIITSACILLALAW	160	
Db	4	KLRTFQSLGYDSSKAYAYAFKVSFNLCIWLSTVI--VAKWGQTSTLANVLASALLGLFW	62	
QY	161	QOQGLWLFHECHQOQTKNRPLNDRTLSLFFGNFLGFSRDWKKDKNTHHAATNVIDHDGD	220	
Db	63	QOQGLAHDELHQQVFOODREWGDLFGAFLGVCVCGFSSWWKKDKNTHHAAPNVHGDDP	122	
QY	221	IDLAPL-----FAFIPG-DLCYKKA SFKAILKIVPYQHLYFTAMLPMLRFSWTG	269	
Db	123	IDTHPLLTWSEHALEMESDVPDEELTRMWSRF-----MVLNQTFWYFPILSFAKLSWCL	176	
QY	270	QSQVWFKEKNOMEYKVYQR--NAFWEQATIVGHWAWFYQLFLPLTWPPLR-VAFYFIISQM	326	



```

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 448 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-366-779--5

Query Match      18.6%  Score 448.5;  DB 1;  Length 448;
Best Local Similarity 27.9%;  Pred. No. 4e-36;
Matches 127;  Conservative 71;  Mismatches 186;  Indels 71;  Gaps 15;

QY  10 LRMKVDGKWLYLSEELVKKHPGGAV-IEQYRNSDATHIFAHFEGSSQAYKOLDLLKKHG 68
    | : | | : | | | | : | | | | : | | | | :
Db  23 LWISIOCK-AVDYSDWKDHPGGFPLKSLAGQEVTDAFVAFHPASTW-----69

QY  69 EHDFELEKLEKRLDKVDINVSAYDVSAQ-----EKKMVSEFEKLRQLKHLDDGLMKANET 124
    | | | | | | | | | | | | | | | | | | | |
Db  70 -----KNLDKFFTGYYLKDYSYSEVSKDYRKLVFEFSKM--GLYD-----KKG 110

QY  125 YFLPKAISTLSIMAFVQLVYG-----WYITSACLLALAAQCGWLTHFCHQOQPT 176
    : | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db  111 HIMP--ATLCFIAMLFAMSYGVFLCEGVLVHLFSGCLMGFLWTQSGWIGHDAGHYMW 167

QY  177 KNRPLNTDISLFFGNFLOGFSRDWKKDKHNTTHHAATNVIDHDGDTDLAP-----LFAF 229
    | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db  168 SDSRLNFMGFIIFAAANCLSGISIGWKWNHNAHACNSLEYDLPDQIYIPFLVWSSKFFGS 227

QY  230 IPGDLCKYKASFKAALKVYPYQHLFTAMLPMLRFSTGQSVQVWFKENOMEKYVYQRN 289
    | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db  228 LTSHFYEKRLTFDSLRSFFYSQHTWTFYIPMCAARLNMYQVSLMLLTKRNVSYRAQE--285

QY  290 AFWEQAATVGHAW-VFYQLFL--LPTWPLRVAVFIISQMGGLLIAHVVTFNHNSVDKY 346
    : | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db  286 -----LLGCLVFSIWPDLVSLCPNNGERIMFVIASLSTVGMQOQV-FSLNHFSSVY 337

QY  347 PANSRILNFAALQILTRNNTPSPFFIDWLWGGGLNYQIEHLLFTMPRCNLNACVKYVKE 406
    : | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db  338 VGRPK-GNNFEKQTDGTLDISCPWMDWPHGGSQFQIEHLLFKPMKPRCNLRKISPYVIE 396

QY  407 WKRENLPYLVDYFDGYAMNLOOLKNWAEHIQAK 441
    | | | | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db  397 LCKKHNLPNYASFSPANEMTLRLRNTA--LQAR 429

RESULT 12
US-08-789-936-5
; Sequence 5, Application US/08789936
; Patent No. 5789220
; GENERAL INFORMATION:
; APPLICANT: Thomas, Terry L.
; APPLICANT: Reddy, Avutu S.
; APPLICANT: Nuccio, Michael
; APPLICANT: Freyssinet, Georges L.
; APPLICANT: Nunberg, Andrew N.
; TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/789,936

```



APPLICANT: THURMOND, JENNIFER  
APPLICANT: CHAUDHARY, SUNITA  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS  
OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: RAE-VENTER LAW GROUP, P.C.  
STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,655  
FILING DATE: 11-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: RAE-VENTER, BARBARA  
REGISTRATION NUMBER: 32,750  
REFERENCE/DOCKET NUMBER: CGNE.124.000US  
TELEPHONE: (650) 328-4400  
TELEFAX: (650) 328-4477  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-834-655-10

Query Match 17.1%; Score 412; DB 2; Length 87;  
Best Local Similarity 90.0%; Pred. No. 1.5e-33;  
Matches 72; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
Qy 364 TRNMTSPFIDWLWGLNYQIEHHLFPTMPCNLNACVYKWKCKENNLPLYLVDYDFG 423  
Db 8 TRNMTSPFIDWLWGLNYQIEHHLFPTMPCNLNRCMKYKWKCAENNLPLYLVDYDFG 67  
Qy 424 YAMNLOQLKNAEHIOAKA 443  
Db 68 YNLNLOQLKNAELVQAKA 87

RESULT 15  
US-08-834-033A-11  
Sequence 11, Application US/08834033A  
Patent No. 6075183  
GENERAL INFORMATION:  
APPLICANT: KNUTZON, DEBORAH  
APPLICANT: MUKERJI, PRADIP  
APPLICANT: HUANG, YUNG-SHENG  
APPLICANT: THURMOND, JENNIFER  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS  
OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LIMBACH AND LIMBACH, L.L.P.  
STREET: 2001 FERRY BUILDING  
CITY: SAN FRANCISCO  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,033A  
FILING DATE: 11-APR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: WARD, MICHAEL R.  
REGISTRATION NUMBER: 38,651  
REFERENCE/DOCKET NUMBER: CGAB-300.USA  
TELEPHONE: (415) 433-4150  
TELEFAX: (415) 433-8716  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-834-033A-11  
Query Match 17.1%; Score 412; DB 3; Length 87;  
Best Local Similarity 90.0%; Pred. No. 1.5e-33;  
Matches 72; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
Qy 364 TRNMTSPFIDWLWGLNYQIEHHLFPTMPCNLNACVYKWKCKENNLPLYLVDYDFG 423  
Db 8 TRNMTSPFIDWLWGLNYQIEHHLFPTMPCNLNRCMKYKWKCAENNLPLYLVDYDFG 67  
Qy 424 YAMNLOQLKNAEHIOAKA 443  
Db 68 YNLNLOQLKNAELVQAKA 87  
Search completed: August 6, 2002, 16:22:16  
Job time: 350 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 01:36:14 ; Search time 81.96 Seconds  
(without alignments)  
4381.603 Million cell updates/sec

Title: US-09-555-093-1  
Perfect score: 1462  
Sequence: 1 gtcacacaaatgctcg.....aataaacctgttttcaa 1462

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	6.6	1617	2	US-08-834-655-1
2	97	6.6	1617	3	US-08-834-033A-1
3	97	6.6	1617	3	US-09-363-574-1
4	68.4	4.7	1684	2	US-08-831-570-1
5	68.4	4.7	1684	2	US-08-831-575-1
6	68.4	4.7	1685	1	US-08-366-779-4
7	68.4	4.7	1685	1	US-08-789-936-4
8	68.4	4.7	1685	4	US-08-934-254-4
9	41.6	2.8	1702	4	US-08-934-254-26
10	41	2.8	7218	1	US-08-232-463-14
11	39.8	2.7	3369	2	US-08-619-198-2
12	39.8	2.7	4170	2	US-08-619-198-2
13	39.8	2.7	5192	2	US-08-619-198-8
14	39.8	2.7	5198	2	US-08-619-198-1
15	39	2.7	2277	1	US-08-676-967-2
16	39	2.7	2277	1	US-08-676-974-2
17	39	2.7	2277	2	US-09-098-487-2
18	37.8	2.6	1129	4	US-09-227-357-40
19	37.6	2.6	144	1	US-08-702-344-26
20	37.2	2.5	19124	2	US-08-487-826B-13
21	36.8	2.5	2223	1	US-08-257-073-4
22	36.4	2.5	719	3	US-08-714-918-11
23	36.4	2.5	719	4	US-09-265-315-11
24	36.4	2.5	719	4	US-09-265-315-11
25	36.4	2.5	719	4	US-09-266-417-11
26	36.4	2.5	1098	3	US-09-248-335-35
27	35.8	2.4	1474	4	US-08-821-994-64

28 35.8 2.4 56516 2 US-08-996-306-1 Sequence 1, Appli  
29 35.8 2.4 56516 4 US-09-338-907-1 Sequence 1, Appli  
30 35.8 2.4 56516 4 US-09-218-207-1 Sequence 1, Appli  
31 35.8 2.4 56520 4 US-09-338-907-179 Sequence 179, App  
32 35.8 2.4 56520 4 US-09-218-207-179 Sequence 179, App  
33 35.2 2.4 991 4 US-08-793-634B-3 Sequence 3, Appli  
34 35.2 2.4 1882 1 US-09-370-253-1 Sequence 1, Appli  
35 35.2 2.4 7218 1 US-08-232-463-14 Sequence 14, Appli  
36 35 2.4 2628 3 US-08-143-219-1 Sequence 1, Appli  
37 35 2.4 2852 3 US-09-027-137-2 Sequence 2, Appli  
38 35 2.4 246240 2 US-08-724-394A-20 Sequence 20, Appli  
39 35 2.4 246240 2 US-08-724-394A-21 Sequence 21, Appli  
40 35 2.4 246240 2 US-08-724-394A-22 Sequence 22, Appli  
41 34.8 2.4 1428 2 US-08-786-999-2 Sequence 2, Appli  
42 34.8 2.4 1428 4 US-08-969-987-8 Sequence 8, Appli  
43 34.8 2.4 1491 4 US-08-999-774A-7 Sequence 7, Appli  
44 34.8 2.4 1494 1 US-08-585-758A-3 Sequence 3, Appli  
45 34.8 2.4 1494 1 US-08-977-818-3 Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
US-08-834-655-1  
; Sequence 1, Application US/08834655  
; Patent No. 5968809  
; GENERAL INFORMATION:  
; APPLICANT: KNUTZON, DEBORAH  
; APPLICANT: MORKERJI, PRADIP  
; APPLICANT: HUANG, YUNG-SHENG  
; APPLICANT: THURMOND, JENNIFER  
; APPLICANT: CHAUDHARY, SUNITA  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS  
; TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: RAE-VENTER LAW GROUP, P.C.  
; STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039  
; CITY: PALO ALTO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/834,655  
; FILING DATE: 11-APR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RAE-VENTER, BARBARA  
; REGISTRATION NUMBER: 32,750  
; REFERENCE/DOCKET NUMBER: CGNE.124.000US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 328-4400  
; TELEFAX: (650) 328-4477  
; TELEX: N/A  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1617 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
US-08-834-655-1

Query Match 6.6%; Score 97; DB 2; Length 1617;  
Best Local Similarity 47.2%; Pred. No. 8.1e-17;  
Matches 415; Conservative 0; Mismatches 440; Indels 24; Gaps 3;

QY	464	tcgcatgtttattagcaacttgcataaggcaacaattcgatcgttggttaaacacatcgaattctgc	523
Db			
Db	536	TGGCTGCGGTCTTTTGGGTCTGTCTTGCCAGCAGTCGGATGGTGTGGCTACGACTTTTTTG	595
QY	524	catacaaccgaaccaagaacagaccttgaatgatactatttcttgtttctttggtaaat	583
Db			
Db	596	CATCACCAAGCTCTTCAGAGACCCTTCTCGGGGTGAATCTTTTCGGGCCCTCTTGGGAGGT	655
QY	584	tctctacaaggattttcaagagatgggtgaaggacaagcataaacctcatcacgctgcc	643
Db			
Db	656	GTCTGCCAGSGCTTCTCGTCCCTGCTGGTGAAGACAAGCACAACTACACACGCCGC	715
QY	644	acaaatgtaattgatcatcagcggtagatcogacttggcaacacttttgcatttatccca	703
Db			
Db	716	CCCAACGCTCACGGCAGAGATCCGACATFAGACCCACCCCTCTGTTGACCTGGAGTGAG	775
QY	704	ggagatttggcaagtataaggccagctttgaaaagcaat-----tctc	748
Db			
Db	776	CATGCGTTGGAGATGTTCTTCGGATGTCCAGATGAGGAGCTGACCCGGCATGTGTCGGGT	835
QY	749	aagattgtaccatatcaacatctctattccaocgcaatgcttccaatgctcgcgtttctca	808
Db			
Db	836	TTCATGTGCTCGAACAGAGACCTGGTTTTACTTCCCATTCTCTCTGTTTGCCCGTCTCTCC	895
QY	809	tggacttggtaagtcagttcfaatgggtatt-----caaagagaatcaaatggagatacaag	862
Db			
Db	896	TGCTGCTCTCCAGTCCAATTCTCTTTGTGTGCCTTAACGGTCAGGCCCCACAACGCCCTCGGGC	955
QY	863	gtctatcaagaanaatgcattcttgggagaagaacaaattgttgacattggcgttgggta	922
Db			
Db	956	GCGCGTGTGCCCATCTCTGTTGGTTCGNAGCAGCTGTCGTTGCGATGCCACTGGACCTGGTAC	1015
QY	923	tctctacaattgttcttattacaacatggccact---tcgggttgcatttcttoattatt	979
Db			
Db	1016	CTGCCACCATGTTCCTGTTTCATCAAGGATCCGTCACATGCTGGTGCTACTTTTTTGGTG	1075
QY	980	tcacaaatgggagaggcccttttgattgctcaagtagtcaacttcaaccataaactctgtt	1039
Db			
Db	1076	TCGCAGCGCGTGTGCGGAACATTGTGTGGCGATCTGTGTTCTCGCTCAACCACAACGGTATG	1135
QY	1040	gataagtatccagccaattctcgaattttaaacacttcgcgccttcaaatittgacc	1099
Db			
Db	1136	CCTGTGATCTCGAAGGAGGCGGTGCGATATGGATTCTTCAGAAGCAGATCATCAG	1195
QY	1100	acacgcaaatgactccatctccattcatattgattggctttggggggactcaattatcaag	1159
Db			
Db	1196	GGTGTGTATCTCCACCCGGGTCTAATTGGCAACTGTTTCAGTGCCTGCCACAACTTTCAAAGATCCAGCGCTGCT	1255
QY	1160	atcgacacacactgttcccacaatgcacgcttgcactcgaatgcttgctggaatat	1219
Db			
Db	1256	ATCGACACCACTGTTTCCCTTCGATGCCTGCCACAACTTTCAAAGATCCAGCGCTGCT	1315
QY	1220	gtgaaagaaatgggtgcaagagaaataatcttcctaccctcgtcgtactctttgacgga	1279
Db			
Db	1316	GTGAGACCTGTGCAAAAAGTACAATGTCCGATACCACACCCCGGTATGATCAGGGA	1375
QY	1280	tatgcaatgaatttgcacaaattgaaaaatatggctgag	1318
Db			
Db	1376	ACTGCAAGCTCTTTTAGCCGTCTGAACAGAGGCTCCAAAG	1414

## RESULT 2

US-08-834-033A-1

; Sequence 1, Application US/08834033A

; Patent No. 6075183

GENERAL INFORMATION:

APPLICANT: KNUTZON, DEBORAH

APPLICANT: MUKERJI, PRADIP

APPLICANT: HUANG, YUNG-SHEN

APPLICANT: THURMOND, JENNIFER

APPLICANT: CHAUDHARY, SUNIT

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS

```

1  TITLE OF INVENTION:  OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
2
3  NUMBER OF SEQUENCES:  26
4  CORRESPONDENCE ADDRESS:
5
6  ADDRESSEE:  LIMBACH AND LIMBACH, L.L.P.
7  STREET:  2001 FERRY BUILDING
8  CITY:  SAN FRANCISCO
9  STATE:  CA
10
11 COUNTRY:  USA
12
13 ZIP:  94111
14
15 COMPUTER READABLE FORM:
16
17 MEDIUM TYPE:  Floppy disk
18
19 COMPUTER:  IBM PC compatible
20
21 OPERATING SYSTEM:  PC-DOS/MS-DOS
22
23 SOFTWARE:  WORDPERFECT 5.1
24
25 CURRENT APPLICATION DATA:
26
27 APPLICATION NUMBER:  US/08/834.033A
28
29 FILING DATE:  11-APR-1997
30
31 CLASSIFICATION:  800
32
33 ATTORNEY/AGENT INFORMATION:
34
35 NAME:  WARD, MICHAEL R.
36
37 REGISTRATION NUMBER:  38,651
38
39 REFERENCE/DOCKET NUMBER:  CGAB-300.USA
40
41 TELECOMMUNICATION INFORMATION:
42
43 TELEPHONE:  (415) 433-4150
44
45 TELEFAX:  (415) 433-8716
46
47 TELEX:  N/A
48
49 INFORMATION FOR SEQ ID NO:  1:
50
51 SEQUENCE CHARACTERISTICS:
52
53 LENGTH:  1617 base pairs
54
55 TYPE:  nucleic acid
56
57 STRANDEDNESS:  single
58
59 TOPOLOGY:  linear
60
61 MOLECULE TYPE:  DNA (genomic)
62
63 US-08-834-033A-1

```

Query Match	Score 97;	DB 3;	Length 1617;
6.68;			

Best Local Similarity 47.2%; Pred. No. 8.1e-17;

Matches	415	Conservative	0	Mismatches	440	Indels	24	Gaps	3
Matches	415	Conservative	0	Mismatches	440	Indels	24	Gaps	3

464	Qy	ctctgcattgttattagcaacttgcattggcaacaattcggatggtttaacaacatgagttctgc	523
465	Qy		
466	Qy		
536	Db	TCGGCTCGCGCTTTTGGGCTCTGTTCTTGGCAGCAAGTGGCGATGTTGGCTCACACATTTTGG	595
537	Db		
538	Db		
524	Qy	catcaacagccaaacaagaacagacaccttgaatgatactattctcttggctctcttggttaat	583
525	Qy		
526	Qy		
596	Db	CATCACCAAGCTCTCCAGGACCGTTTCTGGGGTGATCTTTTCGGCGCCTCTCTTGGAGGT	655
597	Db		
598	Db		
584	Qy	ttcttacaaggaatttctaaagagattggttgaaggacaagcatatacaactcatcagctgcc	643
585	Qy		
586	Qy		
656	Db	GTCTGCCAGGGCTCTCTGCTCTGTGGTGGAGGACAGCAACACTCACCACGCCGCC	715
657	Db		
658	Db		
644	Qy	acaaattgaattgatcatcagctgatatcagacttggaccacttttgcattatttcca	703
645	Qy		
716	Db	CCCAAGTCCACGGCAGGATCCGCACATTGACACCCACCTCTGTTGACCTGGAGTGAG	775
717	Db		
704	Qy	ggagatttgtcaagtataaggccagcttgtgaaaagaact-----tctc	748
705	Qy		
776	Db	CATGCGTTGGAGATGTTCTCGGATGTCCAGATGAGGAGCTGACCCGCATGTGGTGCGGT	835
777	Db		
749	Qy	aagattgaccatatcaaatctctatttcaacgaatgcttccaatgctcggttttcca	808
750	Qy		
836	Db	TTCATGCTCTGACACAGACCTGGTTTTACTTCCCACTTCTCTGTTGCCCGCTCTCC	895
837	Db		
809	Qy	tggactggtcagtcagttcaatgggtatt-----caasagagaatacaaatggagtacaag	862
810	Qy		
896	Db	TGGTGCTCCAGTCCATTCTTTGTGTGCTTACGGCTCAGGCCACAAAGCCCTCGGCG	955
897	Db		
863	Qy	gtctatacaagaagaatgcattctctgggagacaagcaacaattgattggacattgggcttgggta	922
864	Qy		
956	Db	GCGCGTGTGCCCAATCTCGTTTGGTCGAGCAGCTGCTGCTTTCGATGTCACATGGACCTGGTAC	1011
957	Db		
923	Qy	ttctatcaattgttcttattaccacaacatggcgaact---tcggggttgcttatttcattatt	979

Db 1016 CTCGCCACCATGTTCTGTTTCATCAAGGATCCCGTCAACATGCTGGTGTTCTTTTGGTG 1075  
Qy 980 tcaaaaatggaggagggccttttgattgctcaagtagtcaactttcaaccataaactctgtt 1039  
Db 1076 TCGCAGGCGGTGTCGGAACATTTGTCGATGTCGCTCAACCAACACGATG 1135  
Qy 1040 gataagatccagcgaattctcgaattttaaacaactcgcgcctctcaaaattttgacc 1099  
Db 1136 CCGTGATCTCGAAGGAGGCGCTCGATATGGATTCTTCACGAAGCAGATCATCAG 1195  
Qy 1100 acacgaacatgaactccatctccatcattgattggcttgggtggagtcacatcattcag 1159  
Db 1196 GGTCTGATGTCACCGCGGTCTATTGCAACTGGTTCACGGTGGATTGAACATCAG 1255  
Qy 1160 atcgagcaccacttttcccaaatgacacgacgcttgaactcgaatcgaatcgtcgtaaat 1219  
Db 1256 ATCGAGCACCACTTGTCCCTCGATGTCGCCCAACTTTTCAAGATCCAGCGCTGCT 1315  
Qy 1220 gtgaagaatggtgcaagaagataatcttccctaccctcgtcgatgactactttgacgga 1279  
Db 1316 GTCGAGACCTGTGCAAAAGTACAAATGTCGATACCAACACCGGTATGATCGAGGGA 1375  
Qy 1280 tatgcaatgaattgcaacaattgcaaaaatatggctgag 1318  
Db 1376 ACTGCAGAGGCTTTAGCGCTGTGAACGAGGTCTCCAAAG 1414

## RESULT 3

US-09-363-574-1  
; Sequence 1, Application US/09363574  
; Patent No. 6136574  
; GENERAL INFORMATION:  
; APPLICANT: KNUITZON, DEBORAH  
; APPLICANT: MURKERJL, PRADIP  
; APPLICANT: HUANG, YUNG-SHENG  
; APPLICANT: THURMOND, JENNIFER  
; APPLICANT: CHAUDHARY, SUNITA  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS  
; OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: LIMBACH AND LIMBACH L.L.P.  
; STREET: 2001 FERRY BUILDING  
; CITY: SAN FRANCISCO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/363,574  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WARD, MICHAEL R.  
; REGISTRATION NUMBER: 38,651  
; REFERENCE/DOCKET NUMBER: CGAB-202 USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 433-4150  
; TELEFAX: (415) 433-8716  
; TELEX: N/A  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1617 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
US-09-363-574-1

## Query Match

6.6%; Score 97; DB 3; Length 1617;

Best Local Similarity 47.2%; Pred. No. 8.1e-17;

Matches 415; Conservative 0; Mismatches 440; Indels 24; Gaps 3;

Qy 464 tctgcatgtttattagcacttgacgaacaaatctcgatggtgttaacacatgattctgtc 523  
Db 536 TCGGCTGCGCTTTTGGTCTGTTGCGACAGTGGGATGTTGGCTCAGCAGCTTTTGG 595  
Qy 524 catcaacagcgaacaaagaaagacactttgaatgatactattcttcttcttgaataat 583  
Db 596 CATCACCAAGCTTCCAGAGACGTTCTGGGGTGTATCTTTTCGGCGCCCTTTTGGAGGT 655  
Qy 584 tctctacaaggaatttcaagagatggttggaaggaacaaagcaactcaacactcatcacgtcgc 643  
Db 656 GTCTGCCAGGGCTTCTGCTCTGCTGGTGAAGGAGCAACACACATCACCACCGCGCC 715  
Qy 644 acaaatgtaattgatcagcgggtgatcgcacttgggaacacacttttgcacattattcca 703  
Db 716 CCCAACGTCACGCGGAGGATCCGACATTTGACACCCACCTCTGTTGACCTGGAGTGAG 775  
Qy 704 ggagatttgcaagtataaaggccagcgtttgaaaaagaacat-----tctc 748  
Db 776 CATGCGTTGGAGATGTTCTCGGATGTCCAGATGAGGAGCTGACCCGCGCATGTGGTGGCT 835  
Qy 749 aagattgtaccatatcaacatctctatttcacgcgaatgcttccaatgctccgtttctca 808  
Db 836 TTATGCTGCTGAACGAGACCTGGTTTACTTCCCACTTCTCTGTTTCCCGCTCTCTCC 895  
Qy 809 tggactggtcagtcagttcaatgggtatt-----caaagagaatacaaatggagtacaag 862  
Db 896 TGGTGCTCCAGTCCATTCTCTTTGTGCTGCCTAACGGTCAGGCCACACAGCCCTCGGGC 955  
Qy 863 gctatcaaaagaatgcaattctggaagcaagaacaaatggttggacattgggcttgggta 922  
Db 956 GCGCGTGTGCCCATCTCGTTGGTGGTGGAGCAGCTGTGCTTTCGATGCACTTGGACCTGGTAC 1015  
Qy 923 tctctaatgttcttattaccaacatggccact---tcgggttgcttatttcaatt 979  
Db 1016 CTCGCCACCATGTTCTGTTTCATCAAGGATCCCGTCAACATGCTGGTGTTCTTTTGGTG 1075  
Qy 980 tcaaaaatggaggagggccttttgattgctcagtagtcaactttcaaccataaactctgtt 1039  
Db 1076 TCGCAGGCGGTGTCGGAACATTTGTTGGGATCGTGTCTCTGCTCAACCAACAGGTATG 1135  
Qy 1040 gataagtatccagcgaattctcgaattttaaacaacttcgcgcctctcaaaattttgacc 1099  
Db 1136 CCGTGATCTCGAAGGAGGAGCGGTGATATGGATTCTTCACGAAGCAGATCATCAG 1195  
Qy 1100 acacgaacatgaactccatctccatcattgattggcttgggtggagtcacatcattcag 1159  
Db 1196 GGTCTGATGTCACCGCGGTCTATTGCAACTGGTTCACGGTGGATTGAACATCAG 1255  
Qy 1160 atcgagcaccacttgttcccaacaatgcaacgttgcacatctgaaatgcttgcgtgaaatat 1219  
Db 1256 ATCGAGCACCACTTGTCCCTTCGATGCTCGCCCAACTTTTCAAGATCCAGCGCTGCT 1315  
Qy 1220 gtgaagaatggtgcaagaagataatcttccctaccctcgtcgatgactactttgacgga 1279  
Db 1316 GTCGAGACCTGTGCAAAAGTACAAATGTCGATACCAACACCGGTATGATCGAGGGA 1375  
Qy 1280 tatgcaatgaattgcaacaattgcaaaaatatggctgag 1318  
Db 1376 ACTGCAGAGGCTTTAGCGCTGTGAACGAGGTCTCCAAAG 1414

## RESULT 4

US-08-831-570-1  
; Sequence 1, Application US/08831570  
; Patent No. 5959175  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Terry L.

APPLICANT: Nunberg, Andrew N.  
TITLE OF INVENTION: A SUNFLOWER ALBUMIN 5' REGULATORY REGION  
TITLE OF INVENTION: FOR THE MODIFICATION OF PLANT SEED LIPID  
TITLE OF INVENTION: COMPOSITION  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/831,570  
FILING DATE: 09-APR-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 10545  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 743-4366  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1684 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 43..1387  
US-08-831-570-1

Query Match 4.7%; Score 68.4; DB 2; Length 1684;  
Best Local Similarity 44.2%; Pred. NO. 4.le-09;  
Matches 423; Conservative 0; Mismatches 526; Indels 9; Gaps 3;

Qy 381 catattctctgtttaagcgatttcaaacactttcaattatggcatttgcattttatcttc 440  
Db 386 CAACTTTGTGCTTTATAGCAATGCTGTTGCTATGAGTGTATGGGGTTTTGTGTG 445  
Qy 441 agtactctggatgggtatattactcttgcattgtttattagcacttgcattggcaacaattcg 500  
Db 446 AGGGTGTTTGGTACATTTGTTTCTGGTGTGTTGATGGGTCTTTGGATTGAGAGTG 505  
Qy 501 gatggttaacacatgagttctgcatcaacagcccaacaaagacagacctttgaatgata 560  
Db 506 GTTGATTGGACATGATGCTGGCATTATATGGTAGTCTGATTCAAGGCTTAATAAGT 565  
Qy 561 ctattcttctgttctgttgtaattcttcaagagattttcaagagatttgggaagaca 620  
Db 566 TTATGGGTATTTTGCTGCAAAATGCTTTTCAGGAATAAGTATTGGTTGGTGGAAATGGA 625  
Qy 621 agcaaacactcatcacgctgcccaaaatgtaattgatcatgacggtgatcatcgacttgg 680  
Db 626 ACCATAATGCACATCACATTCCTGTATAGCCTTGAATATGACCTGATTTACAATATA 685  
Qy 681 cacaacttttcgcatatttccagagagatttggcaagatataaagccagctttgaaaaag 740  
Db 686 TACCATTCTTGTGTGTCCTTCCAAAGTTTGTGTTGTTTACTACCTCTCATTTCTATGAGA 745  
Qy 741 caattctcaagattgtaccatatacaatctctattttcacgcgaatgcttc---caatgc 797  
Db 746 AAAGTTGACTTTTGACTCTCTTATCAAGATTCTTTGTAAGTTATCAACATGGACATTTT 805

Qy 798 tccgtttctcatggactgggtcagttcagttcaatgggttattcaaaagagaatcaaatggagt 857  
Db 806 ACCCTATTATGTGTGCTGCTAGGCTCAATATGTATGTACAACTCTCTATAATGTTGTGA 865  
Qy 858 acaaggtctatacaaaagaaatgcattctctggagcaagcaaatgttggacattgggett 917  
Db 866 CCAAGAGAAATGTGTCCTATCGAGCTCAGGAACCTCTGGGATGCCCTAGTGTTCGATTT 925  
Qy 918 gggattctatcaaatgttcttattaccacaatgcccacttcgggttgcatttattcatta 977  
Db 926 GGTACCCGCTGCTGTTGTTCTTATGCTTAATTTGGGGTGAAAGAAATATGTTGTTATG 985  
Qy 978 ttccaaaatggagggagggccttttgattgctcacgttagtcaactttcaaccataactcg 1037  
Db 986 CAAGTTTATCAGTGACTGGAATGCAACAAGTTCA---GTTCTCTTGAACCACTTCTCTT 1042  
Qy 1038 tggataagtataccagccaattctctgaatttttaacaacattcgccctcttcaaatgtga 1097  
Db 1043 CAAGTGTGTTTATGTTGGAAGCCT---AAAGGGAATTAATTTGGTTTGAGAAACAACGGATG 1099  
Qy 1098 ccacagcaacatgactccatctccattcattgattggttgggttgactcaattatc 1157  
Db 1100 GGACACTTGCATTTCTTGTCTCTTGGATGGATTGGTTTCATGGGATGCAATTC 1159  
Qy 1158 agatcgagcaccacttgttcccaacaatgccacgttgcaactctgaaatgcttgcgtgaaat 1217  
Db 1160 AAATTGACATCATTTGTTTCCCAAGATGCCCTAGATGCAACCTTAGGAAATCTCGCCT 1219  
Qy 1218 atgtgaagaatggtgcaagaagaataatctcttaoctogtgcatttgaog 1277  
Db 1220 ACGTGATCGAGTTATGCAAGAAACATAATTTGCCCTTACAAATATATGCATCTTTCTCAAAG 1279  
Qy 1278 gatagcaatgaatttgcaacaattgaaaatatgtcgtgagacacattcaagctaaagc 1335  
Db 1280 CCAATGAATGACACTCAGAACATTGGAGAACACAGCATTCGACGCTAGGATATAAC 1337

RESULT 5  
US-08-831-575-1  
; Sequence 1, Application US/08831575  
; Patent No. 597436  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Terry L.  
; APPLICANT: Li, Zhongsen  
; TITLE OF INVENTION: AN OLEOSIN 5' REGULATORY REGION FOR THE  
; TITLE OF INVENTION: MODIFICATION OF PLANT SEED LIPID COMPOSITION  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/831,575  
; FILING DATE: 09-APR-1997  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1684 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 43..1387  
US-08-831-575-1

Query Match 4.7%; Score 68.4; DB 2; Length 1684;  
Best Local Similarity 44.2%; Pred. No. 4.1e-09;  
Matches 423; Conservative 0; Mismatches 526; Indels 9; Gaps 3;

QY 381 catattctctgtttaaacgatttcaacacttccaattatgcatgacattgcatatttatcttc 440  
DB 386 CAACCTTGTGCTTTATAGCAATGCTTTGCTATGAGCTTTATGGGGTTTGTGTTG 445

QY 441 agtatcttgatggtatattactctgcatgtttattagcacttgcacatggaacattcg 500  
DB 446 AGGGTGTGTTGTACATTTGTTTCTGGGTGTTGATGGGTTTCTTGGATTGAGAGTG 505

QY 501 gatgttaacacatgagttctgcatacaacagccaacaaagacagacccttgaatgata 560  
DB 506 GTTGATTGGACATGATGCTGGCCATTATATGTTGCTGATTCAAGGCTTAATAAGT 565

QY 561 ctattctctgttcttggaattcttacaaggattttccaagagatttggtggaaggaca 620  
DB 566 TTATGGGTATTTTCTGCGAATTTGCTTTACAGGAATAAGTATGTTGGTGGGAATGGA 625

QY 621 agcatacaactcatcacgctgcacaaatgtaattgatcatgacggtgatcgcacttgg 680  
DB 626 ACCATAATGCACATCACATTGCTGTAATAGCTTGAATATGACCTGATTTACAATATA 685

QY 681 caeacatttcgcatatttcagagagatttgcgaagtataaagccagccttgaataag 740  
DB 686 TACCATTCTCTTGTGTCTTCCAAAGTTTGTGGTTCACCTCACTCTCATTTCTATGAGA 745

QY 741 caattctcaagatttaccataatcaacatctctatttcaccgcaatgcttc---caatgc 797  
DB 746 AAAGGTTGACTTTTGACCTTTATCAAGATTCCTTTGTAAGTATTAACATGGACATTTT 805

QY 798 tccgtttctcagtgagtgctcagtcagttcaatgggtattcaaaagagaatcaaatggact 857  
DB 806 ACCCTATTATGCTGCTGCTAGGCTCAATATCTATGTACAACTCTCATAAATGTTGTTGA 865

QY 858 acaaggtctatcaagaagaatgcatctctgggagcaagcaacaattgttggacattgggctt 917  
DB 866 CCAAGAGAAATGTGCTATCGAGCTCAGGAACCTCTTGGGATGCTAGTGTCTCGATTT 925

QY 918 gggttattctataaattgttcttattaccaacaacatggccacttcgggtgtcttattcatta 977  
DB 926 GTTACCCGTTGCTTGTCTTGTGCTTGTGCTTAAATGGGGTGAAGAATATGTTGTTATG 985

QY 978 ttccaataaggagagggccttttgattgctcaogtagtcaactttcaaccataactctg 1037  
DB 986 CAAGTTTATCAGTACGACTGGAATGCAACAAAGTTCA---GTTCTCTTGAACCACTTCTCTT 1042

QY 1038 ttgataagtatccagccaattctcgaatttttaacaaacttcgccctcttccaatttga 1097  
DB 1043 CAAGTGTATTATGTTGGAAAGCT---AAAGGGAATAAATTTGTTTGGAAACAAACGGATG 1099

QY 1098 ccacacgcaacatgactccatctccattcattgattgctttgggtgggaactcaattacc 1157  
DB 1100 GGACACTTGACATTTCTTGTCTCTCTTGGATGGATTGGTTTCATGTTGATTGCAATTC 1159

QY 1158 agatcgagcaccaactgtttcccaacaatgccacgttgcaactctgaatgcttgcgtgaaat 1217  
DB 1160 AAATTGAGCATCATTTGTTTCCAGATGCTTAGATGCAACCTTAGGAAAAATCTGCCCT 1219

QY 1218 atgtgaaagaatggtgcaagagaataatcttccctaactctcgatgactactttgacg 1277  
DB 1218 ATGTGAAAGTATTTTGTGCTCAAAATGCTTTTCAGGAATAAGTATGTTGTTGGTGAATGGA 1277

DB 1220 ACGTGATCGAGTTATGCAAGAACAATAATTTGCTTTACAATAATGATCATCTTTCTCAAGG 1279

QY 1278 gatagtcaatgaattgccaacaattgaaaaatatgctgagcacattccaagctaaagc 1335  
DB 1280 CCATGAAATGACACTCAGAACATTTGAGGACACACAGCATTTGCAGCTAGGATATAAC 1337

RESULT 6  
US-08-366-779-4  
; Sequence 4, Application US/08366779  
; Patent No. 5614393  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Terry L.  
; APPLICANT: Reddy, Avutu S.  
; APPLICANT: Nuccio, Michael  
; APPLICANT: Freyssinet, Georges L.  
; APPLICANT: Nunberg, Andrew N.  
; TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A  
; TITLE OF INVENTION: DELTA 6-DESATURASE  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/366,779  
FILING DATE: 30-DEC-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Presser, Leopold  
REGISTRATION NUMBER: 19,827  
REFERENCE/DOCKET NUMBER: 83832YXW  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1685 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-366-779-4

Query Match 4.7%; Score 68.4; DB 1; Length 1685;  
Best Local Similarity 44.2%; Pred. No. 4.1e-09;  
Matches 423; Conservative 0; Mismatches 526; Indels 9; Gaps 3;

QY 381 catattctctgtttaaacgatttcaacacttccaattatgcatgacattgcatatttatcttc 440  
DB 387 CAACCTTGTGCTTTATAGCAATGCTTTGCTATGAGTGTATTTGGGGTTTGTGTTG 446

QY 441 agtatcttgatggtatattactctgcatgtttattagcacttgcacatggaacattcg 500  
DB 447 AGGGTGTGTTGTACATTTGTTTCTGGGTGTTTGGGTTTCTTGGATTGAGAGTG 506

QY 501 gatgtttaaacacatgagttctctgccatcaacagccaacaaagacagacccttgaatgata 560  
DB 507 GTTGATTGGACATGATGCTGGCCATTATATGTTGCTGATTCAAGGCTTAATAAGT 566

QY 561 ctattctctgttcttggtgataattttccaaggattttccaagagattggtggaaggaca 620  
DB 567 TTATGGGTATTTTGTGCTCAAAATGCTTTTCAGGAATAAGTATGTTGTTGGTGAATGGA 626



Db 987 CAAAGTTTATCAGTACTGGAATGCAACAAGTTCA---GTTCTCCTTGAACCACTTCTCTT 1043  
Qy 1038 ttgataagtatccagccaattctcgaattttaacaaacttcgccgctcttcaaaatttga 1097  
Db 1044 CAAAGTGTATGTTGGAAGCCT---AAAGGGAATAAATGGTTTGAGAAACAACCGATG 1100  
Qy 1098 ccacagcaaatgactcatcattcattggtgcttgggtggactcaattatc 1157  
Db 1101 GGACACTTGACATATCTTCTCCTCCTTGGATGGATTGGTTTCATGGTGGATTGCAATTC 1160  
Qy 1158 agatcgagcaccactgttcccaacaatccacggttgaattctgctggaat 1217  
Db 1161 AAATTGAGCATATTTGTTCCAGATGCCCTAGATGCAACCTTAGGAAAATCTCGCCCT 1220  
Qy 1218 atgtgaaagaatgggcaagaagaataatcttcccttacctcgcgatgactattgagc 1277  
Db 1221 ACGTGATCGAGTTATGCAAGAAACATAAATTTGCCITACAAATATGCATCTTCTCAAAG 1280  
Qy 1278 gatagcaaatgttgcaacaaattgaaaaatatgctgagcacattcaagctaaagc 1335  
Db 1281 CCAATGAATGACACTCAGAACATTGAGGAACACAGCATTCGAGGCTAGGGATATAAC 1338

## RESULT 8

US-08-934-254-4

; Sequence 4, Application US/08934254

; Patent No. 6355861

; GENERAL INFORMATION:

; APPLICANT: Thomas, Terry L.

; TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A

; TITLE OF INVENTION: DELTA 6-DESATURASE

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scully, Scott, Murphy &amp; Presser

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/934,254

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Presser, Leopold

; REGISTRATION NUMBER: 19,827

; REFERENCE/DOCKET NUMBER: 83832YXWU

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1685 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: both

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-934-254-4

## Query Match

Best Local Similarity 4.7%; Score 68.4; DB 4; Length 1685;

Matches 423; Conservative 0; Mismatches 526; Indels 9; Gaps 3;

Qy 381 catatttctgtttaagcgatttcaaacatttcaattatgagcattgcatatttcttc 440

Db 387 CAACTTTGTGCTTTTAGCAATGCTGTTTGTCTATGAGTGTTTATGGGGTTTGTG 446

Qy 441 agtatcttgatggtatattacttctgattgtttattagcaacttgcattgccaacattcg 500  
Db 447 AGGGTGTGTTGTACATTTGTTCTGGGTGTTGATGGGTTTCTTTGGATTTCAGAGTG 506  
Qy 501 gatggttaacacatgatttctgccatcaacagccaacaagaacagaccctttgaaatgata 560  
Db 507 GTTGATTGGACATGATGCTGGGCATTATATGTTAGTGTCTGATTCAAGGCTTAATAAGT 566  
Qy 561 ctatttcttcttcttgataatttctacaaggatttccaagagatttccaagagatttgcgaagaca 620  
Db 567 TTATGGGTATTTTGTGTCGAAATTTGCTTCAGGAATAAGTATTTGGTTGGTGAATATGA 626  
Qy 621 agcataacactcatcgcgtgccacaataatgtaattgatcatgacggtgatatcgcattgg 680  
Db 627 ACCATAATGCACATCACATTGCTGTGAATAGCCTTGAATATGACCTGATTACATAATA 686  
Qy 681 caccacttttcgcatttattccaggagatttgcgaagtataagggcagctttgaaaaag 740  
Db 687 TACCATTCTCTGTTGTGCTCTCCAAAGTTTTTTGGTTTCACTCACTCTCATTTCTATGAGA 746  
Qy 741 caattctcaagattgtaccatatcaacatctctatttaccgcaatgcttc---caatgc 797  
Db 747 AAAGTTTGACTTTTGACTCTTTATCAAGATTTCTTTGAAGTTATCAACATTTGGACATTTT 806  
Qy 798 tccgtttctcagactggctcagtcagttcaatgggtatttcaaaagagaatcaaatggagt 857  
Db 807 ACCCTATTATGCTGCTGCTAGGCTCAATATGATGATGATACAACTCTCATAAATGTTGTTGA 866  
Qy 858 acaaggtctatcaaaagaatgcatctctgggagcaagaacaatttggacatttgggctt 917  
Db 867 CCAAGAGAAATGTGCTCTATGAGCTCAGGAACCTTTGGGATGCCCTAGTGTCTTCGATTT 926  
Qy 918 gggatttctcaaatgttcttattaccacaatggccacttcgggttgcatttcttcattc 977  
Db 927 GGTACCCGTTGCTTGTGTTTGTTCCTTAATTTGGGGTGAAGAATATGTTGTTGTTATTG 986  
Qy 978 ttccacaatgggagggccttttggattgctcacgtagctcactttcaaccataactctg 1037  
Db 987 CAAAGTTTATCAGTACTGGAATGCAACAAGTTCA---GTTCTCTTGAACCACTTCTCTT 1043  
Qy 1038 ttgataagtatccagccaattctcgaattttaacaaacttcgccgctcttcaaaatttga 1097  
Db 1044 CAAAGTGTGTTGTTGAAAAGCCT---AAAGGGAATAAATTTGGTTTGAGAAACAACCGATG 1100  
Qy 1098 ccacagcaaatgactcctcatcattgattggttgggttgggagactcaattatc 1157  
Db 1101 GGACACTTGACATTTCTTCTCCTCTTGGATGGATTGGTTTCATGGTGGATTGCAATTC 1160  
Qy 1158 agatcgagcaccactgttcccaacaatgccagttgcaatctggaatgcttgcgtgaaat 1217  
Db 1161 AAATTGAGCATCATTTGTTTCCAGATGCCCTAGATGCAACCTTAGGAAAATCTCGCCCT 1220  
Qy 1218 atgtgaaagaatgggcaagaagaataatcttcccttacctcgcgatgactatttgaag 1277  
Db 1221 ACGTGATCGAGTTATGCAAGAAACATAAATTTGCTTACAAATATATGATCATCTTTCTCCAAG 1280  
Qy 1278 gatagcaaatgttgcaacaaattgaaaaatatggtgagcacattcaagctaaagc 1335  
Db 1281 CCAATGAATGACACTCAGAACATTGAGGAACACAGCATTCGAGGCTAGGGATATAAC 1338

## RESULT 9

US-08-934-254-26

; Sequence 26, Application US/08934254

; Patent No. 6355861

; GENERAL INFORMATION:

; APPLICANT: Thomas, Terry L.

; TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A

; TITLE OF INVENTION: DELTA 6-DESATURASE

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scully, Scott, Murphy &amp; Presser















```
Db 50 WLFSAWFVLPAPVTFPPVRLLCGMVLATAIAAFSPN-----VGH DANHAYSSNPHINRVL 105
Qy 186 SLFFGNFLQGSRDWKKDKHN--THAAATNVIDHGDIDIDLAPFAFIPGDLCKIKASFKA 244
Db 106 GMTY-DFV-GLSFLWRVRYHLYHTYTNILGHDEVIH-----GD-GAVRMSPEQE 153
Qy 245 LKIVPYQHLVFTAMPLMLRSWTSQSVOWFKENQMEYKVKYQNAF--WEQATVIG-HW 301
Db 154 HVGIVRFQOFVITWGLYLFIPFWFLDYVLV--NKGKYHDHKIPFPQPLEASLLGIKL 211
Qy 302 ANWFQVLELLPTWPLRVAYFI-----ISOMGGGLLIAHVTFNH-----NSVDKYP 348
Db 212 LWLGIVFGL----PLALGFSPEVLIGASVTYMTYGVVCTFIMLAHVLESTETPDGE 267
Qy 349 NSRIINNEAALQILTRNN--TPSPFDLWGLGNLYOIEHLLFPTMPRCNLNACVKYKWE 407
Db 268 SGADDEWAICQIRTAFAFNFNFWFCGLNHQVTHLLFPNICHYHPOLENIIDV 327
Qy 408 KKNENLPLVDDYFD-GYAMNLQOLKNMAE 436
Db 328 CQEGVEYKVYPTFKAAIASNRWLEAMGK 357

RESULT 2
FD6C_BRANA STANDARD; PRT; 443 AA.
ID FD6C_BRANA AC P48627;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.99.-).
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94345008; PubMed=8066133;
RA Hitz W.D., Carlson T.J., Booth J.R. Jr., Kinney A.J., Stecca K.L.,
RA Yadav N.S.;
RT "Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA
RT and its expression in a cyanobacterium."
RL Plant Physiol. 105:635-641(1994).
CC -!- FUNCTION: CHLOROPLAST OMEGA-6 FATTY ACID DESATURASE INTRODUCES
CC THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 16:3 AND 18:3 FATTY
CC ACIDS, IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT
CC TO USE FERREDOXIN AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS
CC ESTERIFIED TO GALACTOLIPIDS, SULFOLIPIDS AND PHOSPHATIDYLGLYCEROL.
CC -!- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).
CC -!- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/ OR BE INVOLVED IN METAL ION BINDING.
CC -!- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
FT DOMAIN 202 206 HISTIDINE BOX 2.
FT DOMAIN 362 366 HISTIDINE BOX 3.
SQ SEQUENCE 443 AA; 50755 MW; 27544B33AF2781D8 CRC64;

Query Match 7.2%; Score 174.5; DB 1; Length 443;
Best Local Similarity 23.7%; Pred. No. 2.1e-06;
Matches 96; Conservative 75; Mismatches 139; Indels 95; Gaps 26;

Qy 57 AYKOLDLLKKEHDEFELEKLEKR---LDKVDINVSAYDVSVAQEK-KMVES--FEKLR 110
Db 36 AVRPIDLLKLT-GTRTEFL-VPAAKRIGCIKAVFPVPAPPSADNAEDRQLAESYGFKOIG 93
Qy 111 QKLHDDGLMK-----ANETFL-----FKAISTLSIMAF-----FYQLYLGWYITSACLL 156
Db 94 QDLDPNVTLKIDMTLPKEVEIDDVRAKWSV-LISVTSYALGLFMIAKAPWY-----LL 147
Qy 157 ALAWQOFGW-----LTHEFCHOQPTKRNPLNDTI-SLFFGNFLQGSRDWKKDKHNTH 208
Db 148 PLAWAWTGTAVTGFEVIGHDCAHKSFSEKNKLVEDIVGTALPLV--YPPEWRFKHDRH 205
Qy 209 HAATNVIDHD-----GDIDLAPLPAFIPGDLCKYKASEKAILKIVPQHLVFTAML 260
Db 206 HAKTNMLVHDTPAQVPPEEFDSSPV-----LRKAIT-----FGYGPTR 244
Qy 261 PMLRFS-WTGQSVQVFKENQMEYKVKYQRN-----AFWEQATIVGHAWVYQLF 309
Db 245 PWSLTAHW---VNMHF--NLKFRPSEVRVKISLACVFAF---MAVG-WPLLIYKVG 293
Qy 310 LLPTWPLRVAVFIISQMGGLLIAHVTFNHSVDKYPANSRILNNAALQILTRNNWTP 369
Db 294 VL-GW---VKFWLMPWLGJYHFWMSS-TFTMVHTAPHIFPKPADENNAQAOLNGTVHCDY 348
Qy 370 SPFTDLWGLGNLYOIEHLLFPTMPRCNLNACVKYKKE-WCKENNL 413
Db 349 PSWIEILCHIDINVHPHISPRIPSYNLRAAHQSIQENWGWKTYNL 393

RESULT 3
FD6C_ARATH STANDARD; PRT; 448 AA.
ID FD6C_ARATH AC P46312;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.99.-).
GN PAD6 OR FADC OR AT4G30950 OR F6118.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE OF 1-418 FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=95148736; PubMed=7846158;
RA Falcone D.L., Gibson S., Lemieux B., Somerville C.R.;
RT "Identification of a gene that complements an Arabidopsis mutant
RT deficient in chloroplast omega 6 desaturase activity."
RN Plant Physiol. 106:1453-1459(1994).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansorge W., Brandt P., Grivell L., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Bontury M., Bancroft I.,
RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
```

RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,  
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,  
RA Berneser S., Hempel S., Feldpausch M., Lamberth S., van den Daele H.,  
RA De Keyser A., Buyshaert C., Gielen J., Villarroel R., De Clercq R.,  
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
RA Clark L., Doggett J., Hall S., Kay M., Lennard-N., McIlroy K., Mayes R.,  
RA Pettett A., Rajadream M.A., Lyne M., Benes V., Reckmann S.,  
RA Borkova D., Bloecker H., Scharf M., Grimm M., Loehner T.-H.,  
RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,  
RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,  
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,  
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
RA Chedors F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,  
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,  
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
RA Sehkon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,  
RA Latreille P., Courtnay L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,  
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,  
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,  
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
RA Granat S., Shohdy N., Hasegawa A., Rodriguez M., Hoffman J., Till S.,  
RA Chen E., Marra M., Martienssen R., McCombie W.R.,  
RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
RA thaliana";  
RA Nature 402:769-777(1999).  
CC -1- FUNCTION: CHLOROPLAST OMEGA-6 FATTY ACID DESATURASE INTRODUCES  
CC THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 16:3 AND 18:3 FATTY  
CC ACIDS, IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT  
CC TO USE FERREDOXIN AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS  
CC ESTERIFIED TO GALACTOLIPIDS, SULFOLIPIDS AND PHOSPHATIDYLGLYCEROL.  
CC -1- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).  
CC -1- DEVELOPMENTAL STAGE: HIGHEST LEVELS FOUND IN EXPANDING LEAVES.  
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE  
CC AND/ OR BE INVOLVED IN METAL ION BINDING.  
CC -1- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U09503; AAA2800.1; -;  
CC EMBL: AL022198; CAA18198.1; -;  
CC InterPro: IPR001225; FA\_desaturase.  
CC Pfam: PF00487; FA\_desaturase; 1.  
CC ProDom: PD001081; FA\_desaturase; 1.  
CC KX Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;  
CC Transit peptide.  
CC TRANSIT 1 69 CHLOROPLAST (BY SIMILARITY).  
CC CHAIN 70 448 OMEGA-6 FATTY ACID DESATURASE.  
CC DOMAIN 171 175 HISTIDINE BOX 1.  
CC FT DOMAIN 207 211 HISTIDINE BOX 2.  
CC FT DOMAIN 367 371 HISTIDINE BOX 3.  
CC FT SEQUENCE 448 AA; 51225 MW; C3AC72FB28BFB287 CRC64;

Query Match

6.9%; Score 165.5; DB 1; Length 448;

Best Local Similarity 23.7%; Pred. No. 1.le-05;  
Matches 82; Conservative 60; Mismatches 133; Indels 71; Gaps 19;  
QY 88 NVSAVDVSAOEKQWSEFEKRLQKLDGLMKLAKNETYFLFKALISTLSIMAFAYLQYL 147  
DB 104 NVTLDKIDMTLPKEVEI-----DD--LKALKSVLISVTSYTLGL-----FMIKSP 148  
QY 148 WYITSACILALAWQOPGW-----LTHEFCHQOQTKRNLNDTI-SLFFGNFLOGFSRD 199  
DB 149 WY-----LLPLAWTGTATGTFVIGHDCAKSKNKLVEDIVGTTLAEPLV--YFYE 201  
QY 200 WVKDKHNTHTAAATVVIDHGDIDLAPLFAFIPGDLCKYKASEFEKAILKIVPQHLYFTAM 259  
DB 202 PWRFKDRHHRHAKTNMLVHD-----TAMQPPVPEEFESSPMVRKAIL-----FGYGP 248  
QY 260 LPMILRES-WTQSGVQWVFKENOMEYKVKYQNR-----AFWEQATTVGHWAWEYQL 308  
DB 249 RPWLISIAHW-----VNVHFE--NLKFRASEVNRVKISLACVFAF-----NAVGVWPLIVTKV 297  
QY 309 FLPTWPLRVAYFIISOMGGGLLHAHVVTFNHNSVDKYPANSRILNFAALQILTRNMT 368  
DB 298 GIL-GW---VKFELMPWLGYHFHWS-TFTMVHTHTAPHIPKPADEWNAQAQLNGTVHCD 352  
QY 369 PSPFDLWGLNGLQIEHHLFPTMPCNLNACVKYVKE-WCKENML 413  
DB 353 YPSWIEILCHDINHPIHSPRIPSYNLRAAHESIQENWGKYTNL 398  
RESULT 4  
FD6C\_SP10L STANDARD; PRT; 447 AA.  
ID FD6C\_SP10L STANDARD; PRT; 447 AA.  
AC P48629;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.99.-).  
GN FAD6.  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 66-78.  
RC STRAIN=CV, SUBITO; TISSUE=Leaf;  
RX MEDLINE=95036044; PubMed=7948918;  
RA Schmidt H., Dresselhaus T., Buck F., Heinz E.;  
RT "Purification and PCR-based cDNA cloning of a plastidial n-6  
RT desaturase";  
RL Plant Mol. Biol. 26:631-642(1994).  
CC -1- FUNCTION: CHLOROPLAST OMEGA-6 FATTY ACID DESATURASE INTRODUCES  
CC THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 16:3 AND 18:3 FATTY  
CC ACIDS, IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT  
CC TO USE FERREDOXIN AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS  
CC ESTERIFIED TO GALACTOLIPIDS, SULFOLIPIDS AND PHOSPHATIDYLGLYCEROL.  
CC -1- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).  
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE  
CC AND/ OR BE INVOLVED IN METAL ION BINDING.  
CC -1- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X78311; CAA55121.1; -;  
CC InterPro: IPR001225; FA\_desaturase.  
CC Pfam: PF00487; FA\_desaturase; 1.  
CC ProDom: PD001081; FA\_desaturase; 1.

```
KW Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
KW Transit peptide.
FT TRANSIT 1 65 CHLOROPLAST.
FT CHAIN 66 447 OMEGA-6 FATTY ACID DESATURASE.
FT DOMAIN 171 175 HISTIDINE BOX 1.
FT DOMAIN 207 211 HISTIDINE BOX 2.
FT DOMAIN 367 371 HISTIDINE BOX 3.
SQ SEQUENCE 447 AA; 51306 MW; 2BA7C87FFF95350E CRC64;

Query Match 6.68; Score 158.5; DB 1; Length 447;
Best Local Similarity 20.88; Pred. No. 3.9e-05;
Matches 82; Conservative 60; Mismatches 126; Indels 127; Gaps 18;

Qy 89 VSAVDVSAQE-----KKMVE--SFEKLROKLHDDGLMK-----AN 122
Db 63 VOAVAVPAQSPAPPTDTEHLQALARYGFOIGEPDPDVTMRDIITSLPKQVFEIN 122
Qy 123 EYFELFKRAIST--LSIMAF---FYLOYLGYIITSACLLALAWQOFGW-----LTHEF 170
Db 123 DT----KAWGTVLISVTSYALGIFMIKAPWY-----LLPLAWAWTGTAITGFEVIGHDC 173
Qy 171 CHQOPTNRPLNDI-SLFFGNFLQGSRDWKKHNTHTHAATVVDHGDIDILAPLAF 229
Db 174 AHKFSKNKLVEDIVGTIAFMPLI--YPYEPWREKHDQHTKTNLRED-----TAW 223
Qy 230 IPGDLCKYKASFKAALKIPYQHLFTAMLPMLRFSWTGOSVOWFKEKQMKYVQRN 289
Db 224 LP--IMKEDISSGLRKALIYAY-----GPLRTW----- 252
Qy 290 AFWEQATIVGHMAWVFYQL-----FLLPTWPL-----RVAY 320
Db 253 -----SIAHMLKVFHFNKDRQSEKVRATISLAFAFVIGWPLIYKTVIGWIKF 305
Qy 321 FIISQMGGLLIAHVFNHNSVDKYPANSRLNFAALQILTRNTPTSPFIDWLNGGL 380
Db 306 WLMPLGYHFHMS--TFTIVHTAPHIPFKSKKEWNAQAQLSGTVHCDYPRWIELCHDI 364
Qy 381 NYQTEHLFTMPRCNLNACVYVKE--WCKENLNP 414
Db 365 SVHPIPHISPISYNLRAAQNLSNENWGEVNLKP 399

RESULT 5
FD6C SOYBN STANDARD; PRT; 424 AA.
AC P48628;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
OS Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.99.-).
OC Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OC NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94345008; PubMed=80661633;
RA Hitz W.D., Carlson T.J., Booth J.R., Kinney A.J., Stecca K.L.,
RA Yadav N.S.;
RT "Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA
RT and its expression in a cyanobacterium.";
RL Plant Physiol. 105:635-641(1994).
CC -!- FUNCTION: CHLOROPLAST OMEGA-6 FATTY ACID DESATURASE INTRODUCES
CC THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 16:3 AND 18:3 FATTY
CC ACIDS, IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT
CC TO USE FERREDOXIN AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS
CC ESTERIFIED TO GALACTOLIPIDS, SULFOLIPIDS AND PHOSPHATIDYLGLYCEROL.
CC -!- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).
CC -!- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
```





CC ESTERIFIED TO PHOSPHATIDYLCHOLINE AND, POSSIBLY, OTHER  
CC PHOSPHOLIPIDS.  
CC -1- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum.  
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE  
CC AND/ OR BE INVOLVED IN METAL ION BINDING.  
CC -1- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L26296; AAA32782.1; -;  
DR EMBL; AP002063; BAB01960.1; -;  
DR EMBL; AC069473; AAC51042.1; -;  
DR InterPro: IPR001225; FA\_desaturase.  
DR Pfam: PF00487; FA\_desaturase; 2.  
DR ProDom: PD001081; FA\_desaturase; 1.  
KW Oxidoreductase; Fatty acid biosynthesis; Endoplasmic reticulum;  
Transmembrane.  
FT TRANSMEM 56 76 POTENTIAL.  
FT TRANSMEM 117 137 POTENTIAL.  
FT TRANSMEM 179 199 POTENTIAL.  
FT TRANSMEM 225 243 POTENTIAL.  
FT TRANSMEM 252 272 POTENTIAL.  
FT DOMAIN 105 109 HISTIDINE BOX 1.  
FT DOMAIN 141 145 HISTIDINE BOX 2.  
FT DOMAIN 315 319 HISTIDINE BOX 3.  
FT SEQUENCE 383 AA; 44047 MW; 8815ADD2D3BBC982 CRC64;

Query Match 5.8%; Score 140; DB 1; Length 383;  
Best Local Similarity 21.5%; Pred. No. 0.00091;  
Matches 78; Conservative 35; Mismatches 100; Indels 150; Gaps 19;  
QY 136 IMAFAFY-----LQYLGW---YITSACLLALAWQFGWLTHEFCHQOPTKNR 179  
DB 61 IIASCFYVATNTFSLLPQPLSYLAWPLYWACQGVLTGIWV-----IAHECGHHAFFSDYQ 116  
QY 180 PLNDTISLFFGNFL--QGFSRDWMDKHNTHAATNVIDHDGIDIDLAPLFAFIPGDLCKY 237  
DB 117 WLDDTVGLIFHSFLVYPFS---WKYSHRRHSHNTGSLERDE-----VFVP---KQ 161  
QY 238 KASFEKAILKIVPYQHLYFTAMLPLRFSWTGQSVQVFKENQMEYKVKYORNAFWEOATI 297  
DB 162 KS-----AIKW-----YKGLNLP----- 175  
QY 298 VGHWAWVFYQLFLLPTWPLRVAVFIISOMGGGL---LIAHVVTFN-HNSVDKYPANSRI- 352  
DB 176 LGRIMMLTVQFVL--GMPLYLAFNVSGRPYFACHFFPNAPYNDRERLOIYLSADAGIL 233  
QY 353 -----LNNFAALQ-----ILTRNMTSPSFI-----DWLWGG 379  
DB 234 AVCFGLYRYAAQGMASMICLYGVPLLIVNAFLVLYLQHTHPSLPHYDSSEMDWLGA 293  
QY 380 L-----NQTEHHLFPPTMPCNLCNACVKYKWKCKENNLVYLDVY - 420  
DB 294 LATVDRDYGILNKVFHNITDTHVAHLFLSTMPHYNAMEATKAIK-----PILGDYQ 345  
QY 421 FDG 423  
DB 346 FDG 348  
RESULT 8  
FD6E\_BRAJU  
ID FD6E\_BRAJU STANDARD; PRT; 384 AA.  
AC Q39287;  
DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Omega-6 fatty acid desaturase, endoplasmic reticulum (EC 1.14.99.-)  
DE (Delta-12 desaturase).  
OS Brassica juncea (leaf mustard) (Indian mustard).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3707;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. 651-2-5-7-4;  
RA Singh S.P., van der Heide T., McKinney S., Green A.;  
RT "Nucleotide sequence of a cDNA from Brassica juncea encoding a  
RT microsomal omega-6 desaturase";  
RL (in) Plant Gene Register PGR95-107.  
CC -1- FUNCTION: ER (MICROSOMAL) OMEGA-6 FATTY ACID DESATURASE INTRODUCES  
CC THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 18:3 FATTY ACIDS,  
CC IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT TO USE  
CC CYTOCHROME B5 AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS  
CC ESTERIFIED TO PHOSPHATIDYLCHOLINE AND, POSSIBLY, OTHER  
CC PHOSPHOLIPIDS (BY SIMILARITY).  
CC -1- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum.  
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE  
CC AND/ OR BE INVOLVED IN METAL ION BINDING.  
CC -1- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X91139; CAAG2578.1; -;  
DR InterPro: IPR001225; FA\_desaturase.  
DR Pfam: PF00487; FA\_desaturase; 2.  
DR ProDom: PD001081; FA\_desaturase; 1.  
KW Oxidoreductase; Fatty acid biosynthesis; Endoplasmic reticulum;  
Transmembrane.  
FT TRANSMEM 56 76 POTENTIAL.  
FT TRANSMEM 84 104 POTENTIAL.  
FT TRANSMEM 117 137 POTENTIAL.  
FT TRANSMEM 180 200 POTENTIAL.  
FT TRANSMEM 226 246 POTENTIAL.  
FT TRANSMEM 253 273 POTENTIAL.  
FT DOMAIN 105 109 HISTIDINE BOX 1.  
FT DOMAIN 141 145 HISTIDINE BOX 2.  
FT DOMAIN 316 320 HISTIDINE BOX 3.  
FT SEQUENCE 384 AA; 44315 MW; 43AF7CE9861A492 CRC64;  
Query Match 5.7%; Score 138.5; DB 1; Length 384;  
Best Local Similarity 22.4%; Pred. No. 0.0012;  
Matches 78; Conservative 37; Mismatches 114; Indels 119; Gaps 17;  
QY 136 IMAFAFY-----LQYLGWITSAC---LLALAWQFGWLTHEFCHQOPTKNR 179  
DB 61 IVASCFYVATTFYPLPLPPLSYLAWPLYWACQGVLTGIWV-----IAHECGHHAFFSDYQ 116  
QY 180 PLNDTISLFFGNFL--QGFSRDWMDKHNTHAATNVIDHDGIDIDLAPLFAFIP---GDL 234  
DB 117 WLDDTVGLIFHSFLVYPFS---WKYSHRRHSHNTGSLERDE-----VFVPKKKSDI 165  
QY 235 KYKASFEKAILKIVPYQHLYFTAMLPLRFSWT-GQSVQVFKENQMEYKVKYORNAFW 293  
DB 166 KWKYKGLNLPGRV-----MLTVQFTLGNPLYWAFNVSGRPYEGFACHFHP 213  
QY 294 QATIVGHWAVFYQLFLLPTWPLRVAV-----FIISOMGGGLIAH-----VVT 337  
DB 214 NAPIYNDRERL--QIVVSDAGILAVCYGLYRYAAAGVASMVCVLYGVPILLIVNAFLVLI 271

```

QY 338 F---NHNSVDKYPANSRLNNFAALQILTRNMTSPFIDWLWGL----- 380
D  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 272 YLQHTPSLPHYSSE-----WDLRGALATVDRDYGLNLKVF 309
QY 381 ----NYOIEHHLFTMPRCNLACVKYKWKCKENNLPLYLVDDY-FDG 423
D  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 310 HNTDTHVAHHLFTMPHYHAMEVTRAIK-----PILGDYQFDG 349

RESULT 9
D12_CREAL
ID D12_CREAL STANDARD; PRT; 375 AA.
AC O81931;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Delta(12) fatty acid dehydrogenase (EC 1.14.99.33) (Crepenynate
DE synthase) (Delta-12 fatty acid acetylase).
OS Crepis alpina.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Crepis.
OX NCBI_TaxID=72610;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98239771; PubMed=9572738;
RA Lee M., Lennan M., Banas A., Bafor M., Singh S., Schweizer M.,
RA Nilsson R., Liljenberg C., Dahlqvist A., Gummeson P.O., Sjoedahl S.,
RA Green A., Stymne S.;
RT "Identification of non-heme diiron proteins that catalyze triple bond
RT and epoxy group formation."
RL Science 280:915-918(1998).
CC -1- FUNCTION: CHANGES THE DELTA-12 DOUBLE BOND OF LINOLEIC ACID INTO A
CC TRIPLE BOND IN THE BIOSYNTHESIS OF CREPENYNIC ACID.
CC -1- CATALYTIC ACTIVITY: Linoleate + AH(2) + O(2) -> crepenynate + A +
CC H(2)O.
CC -1- COFACTOR: Iron.
CC -1- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: SEED.
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE AND/
CC OR BE INVOLVED IN METAL ION BINDING.
CC -1- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y16285; CAA76158.2;
CC InterPro: IPR001225; FA_desaturase.
CC Pfam: PF00487; FA_desaturase; 1.
CC ProDom: PD001081; FA_desaturase; 1.
CC Oxidoreductase; Fatty acid biosynthesis; Transmembrane; Iron.
CC TRANSMEM 54 74 POTENTIAL.
CC TRANSMEM 77 97 POTENTIAL.
CC TRANSMEM 110 130 POTENTIAL.
CC TRANSMEM 172 192 POTENTIAL.
CC TRANSMEM 218 238 POTENTIAL.
CC TRANSMEM 242 262 POTENTIAL.
CC TRANSMEM 98 102 HISTIDINE BOX 1.
CC DOMAIN 134 138 HISTIDINE BOX 2.
CC DOMAIN 308 312 HISTIDINE BOX 3.
CC SEQUENCE 375 AA; 43382 MW; B31F5A64DD0C2926 CRC64;

```

Query Match 4.88; Score 115.5; DB 1; Length 375;  
 Best Local Similarity 20.4%; Pred. No. 0.074; Indels 133; Gaps 17;  
 Matches 67; Conservative 39; Mismatches 90;

```

QY 135 SIMAFAY-----LOYLGM---YITSACLALAWQFGWLTHFEC HQPKN 178
D  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 53 AIIAYIFVLADKYIPILPAPLAVLAWPLYWFCQASILTGLW-----IGHECGHAFSDY 108
QY 179 RPLNDTISLFGNFLOG--FSRDWKKDKHNTHTHAATNVIDHGDIDDLAPLFAFIPGDLCK 236
D  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 109 QWVDVTGVIILHSFLMTPYFS---WKYSRHHNHNATNSLDNDE-----VYIPKS--K 155
QY 237 YKASFEKAILKIVP----YQHLFTFAMLPMLRFESWVGQVWFKENOMEKYKVYORNA-- 290
D  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 156 AKVALIYKVLNHPGRLIMFTITLGFPLYLFT-----NISKKYERFANH 202
QY 291 -----FMEQATIVGHAWWVYQLFLPTWPLRVAYFI---ISOMG----- 327
D  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 203 FDPMSPIFKER-----FOVLLSDGLLAVLYGKLAVALAKGAAWVTCIYGIPVL 253
QY 328 GGLLIAHVVTENHN--SVDKYPANSRLNNFAALQILTRNMTSPFIDWLW----- 377
D  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 254 GVFFFDIITLYLHHTLHSLPHYDSS-----EWNMLRGALST 289
QY 378 -----GGLN-----YQIEHHLFPPTMP 393
D  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 290 IDREGFLNSVLHDVTHTHVMHLESYIP 318

RESULT 10
SYGB_THEME
ID SYGB_THEME STANDARD; PRT; 672 AA.
AC Q9WY60;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycyl-tRNA synthetase beta chain (EC 6.1.1.14) (Glycine--tRNA ligase
DE beta chain) (GlyRS).
GN GLYS OR TM0217.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Uitterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RA "Evidence for lateral gene transfer between Archaea and Bacteria from
RA genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
CC -1- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) -> AMP + diphosphate
CC + glycyl-tRNA(Gly).
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains
CC (by similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE001706; AAD35309.1;
CC TIGR: TM0217;
CC InterPro: IPR002311; tRNA_synth_2f.
CC Pfam: PF02092; tRNA_synth_2f; 1.
CC PRINTS: PR01045; TRNASYNTHGB.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

```

```
KW Complete proteome.
SQ SEQUENCE 672 AA; 77813 MW; 612B402052DD429A CRC64;

Query Match 4.68; Score 112; DB 1; Length 672;
Best Local Similarity 18.4%; Pred. No. 0.29;
Matches 94; Conservative 65; Mismatches 149; Indels 204; Gaps 22;

QY 15 DGKWLYLEEL- ---VKHGPCGAVTEQYRNSDATHIFHAFEGSS- ---QAYKOLDLL 64
DB 18 KYLELEPEELIYAVKHHQHSFAHK- ---GTLTNTFAVODGPPNNVVGKGV- IN 313
QY 65 KKHGEHDFLEKLEKRLDKVDINVSADVSVAQEK- ---KMVESFEKLRQKLHDDGLM 119
DB 314 ARLEDARYYFKDLETPLEKKNEXLKEI- ---VFQKLGTLVDKVERIKKISQRLCEDLKL 370
QY 120 KANETIFYLFKATSLTSMFAFYLOYLGWYITSACLLALAWQOFGWLTHFCHQOQPKNR 179
DB 371 PGSEFTQVLEA- ---ASICKADIASK- ---VVEPEPELQGVNGR 406
QY 180 - - - - -PLNDTISLFFGNFLOGFSRDWKKDKHNT 207
DB 407 IYALREGINEINAEIDHYSEEPOTVIGSILGIADRIDTIVGNFAGNVPTSSKDPYGL 466
QY 208 HHAATNVI- ---DHDGDDIDLAPLFAFPGDLCKYKASPEKAILKIVPYQHLYFTAMLPLM 263
DB 467 KSKADTIFRIIRKNEWDISLEELLTF- ASSLVGYRLSEE- ---L 505
QY 264 RFSWTGQSVQVFNQMEYKQYOR- NAFWEQATIVGHAWVFFQLFLLPTWPLRVAYFI 322
DB 506 ETFFAGREYQFLNELGISFSDVARAVNHLWK- ---PLR- --- 540
QY 323 ISQMGGLLIA- ---HVVTFNHNSVD- ---KYPANSRLN 354
DB 541 - - - - -GILSRAEQEISEKPEFODLFVGFPERVNIKNHDSFTKFDGALFKEEKKLMN 594
QY 355 NFAALQILTRNMTPTSPFDLWGLNLYQIEHILFPTMPRONLACVYKVEKWKNNLP 414
DB 595 KF- ---YEVKEKVLKALERLNYREALQYLIET- ---KP 624
QY 415 YLVDDYFDG- YAM- ---NLQOLKNMAE 436
DB 625 Y- IDEYFDNVFVMVKRDLRLVNRSLFLKNIDE 655

RESULT 11
FD61_SOYBN
ID FD61_SOYBN STANDARD; PRT; 387 AA.
AC P48630;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Omega-6 fatty acid desaturase, endoplasmic reticulum isozyme 1
DE (EC 1.14.99.-).
GN FAD2-1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=96151506; PubMed=8587990;
RA Heppard E.P., Kinney A.J., Stecca K.L., Miao G.H.;
RT "Developmental and growth temperature regulation of two different
RT microsomal omega-6 desaturase genes in soybeans.";
RL Plant Physiol. 110:311-319(1996).
CC -1- FUNCTION: ER (MICROSOMAL) OMEGA-6 FATTY ACID DESATURASE INTRODUCES
CC THE SECOND DOUBLBOND IN THE BIOSYNTHESIS OF 18:3 FATTY ACIDS.
CC IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT TO USE
CC CYTOCHROME B5 AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS
CC ESTERIFIED TO PHOSPHATIDYLCHOLINE AND, POSSIBLY, OTHER
```

CC Bacteria; Cyanobacteria; Oscillatoriales; Spirulina.  
OX NCBI\_TaxID=1156;  
RN  
RP SEQUENCE FROM N.A.  
RA Murata N., Deshnlum P., Tasaka Y.;  
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CAN INTRODUCE A SECOND CIS DOUBLE BOND AT THE DELTA 12  
CC POSITION OF FATTY ACID BOUND TO MEMBRANES GLYCEROLIPIDS. THIS  
CC ENZYME IS INVOLVED IN CHILLING TOLERANCE BECAUSE THE PHASE  
CC TRANSITION TEMPERATURE OF LIPIDS OF CELLULAR MEMBRANES DEPENDS ON  
CC THE DEGREE OF UNSATURATION OF FATTY ACIDS OF THE MEMBRANE LIPIDS.  
CC -1- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X86736; CAA60415.1; -;  
DR InterPro: IPR001225; FA\_desaturase.  
DR Pfam: PF00487; FA\_desaturase; 1.  
DR ProDom: PD001081; FA\_desaturase; 1.  
DR PROSITE; PS00574; FATTY\_ACID\_DESATUR\_2; FALSE\_NEG.  
KW Oxidoreductase; Fatty acid biosynthesis; Membrane.  
FT DOMAIN 89 93 HISTIDINE BOX 1.  
FT DOMAIN 125 129 HISTIDINE BOX 2.  
FT DOMAIN 286 290 HISTIDINE BOX 3.  
SQ SEQUENCE 351 AA; 40928 MW; 125A9F1E07E5EE97 CRC64;  
  
Query Match 4.6%; Score 110.5; DB 1; Length 351;  
Best Local Similarity 21.3%; Pred. No. 0.17;  
Matches 64; Conservative 43; Mismatches 121; Indels 73; Gaps 11;  
  
Qy 128 FKAISTLSIMAFYQVLYGWYTSACLLALAW-----QQFGWLTHEFCHQOQTKNRP 180  
Db 42 YKAMTNTIINVMVGLGWIATAPWELPVVFTGTALTGFVIGHDCGRSFSRNWV 101  
Qy 181 LNDTIS-LFFGNFLOGESRDWKKHNTHTAATNVIDHDG-----DIDLAPLFAFIP 231  
Db 102 VNDWGHILFLPIIPFHS--WRIGHNHHKYTNRMELDNLAWQWPKRKEEQNAGKFMQVT 159  
Qy 232 GDLCYKASFEKAILKIVPYQHLVFTAMPLMLRFSWTSQSVQWV---FKENOMEXKVYQR 288  
Db 160 YDLFRGRA-----WWIGSILHWSIHFDWTYKFEKQKQQ 193  
Qy 289 NAFWEQATVGHWA-----WVFYQLFLPTWPLRVAYFIISOMGGGLLIAHV 336  
Db 194 VKF-SLLVIGAAAFPTMTLTIGVWGFKVFWIP-W-LVFHEWMS-----TF 239  
Qy 337 TENHNSVDKYKANSRLNFAALQILITRNMTSPFDIWLWGLNLYQIEHLFPTMPRCN 396  
Db 240 TLLHHTIADIPPEQWHEAESQLSGTVHCNYSRWGEFLCHDINVPHIPHHVTTAIPYN 299  
Qy 397 L 397  
Db 300 L 300  
  
RESULT 13  
FD62\_SOYBN  
ID FD62\_SOYBN STANDARD; PRT; 383 AA.  
AC P48631;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Omega-6 fatty acid desaturase, endoplasmic reticulum isozyme 2  
DE (EC 1.14.99.-).  
DE FAD2-2.  
OS Glycine max (Soybean).

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
CC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
OX NCBI\_TaxID=3847;  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Epicotyl;  
RX MEDLINE=96151506; PubMed=8587990;  
RA Heppard E.P., Kinney A.J., Stecca K.L., Miao G.H.;  
RT "Developmental and growth temperature regulation of two different  
RT microosomal omega-6 desaturase genes in soybeans";  
RL Plant Physiol. 110:311-319(1996)  
CC -1- FUNCTION: ER (MICROSOMAL) OMEGA-6 FATTY ACID DESATURASE INTRODUCES  
CC THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 18:3 FATTY ACIDS,  
CC IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT TO USE  
CC CYTOCHROME B5 AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS  
CC ESTERIFIED TO PHOSPHATIDYLCHOLINE AND, POSSIBLY, OTHER  
CC PHOSPHOLIPIDS.  
CC -1- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum.  
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE  
CC AND/ OR BE INVOLVED IN METAL ION BINDING.  
CC -1- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L43921; AAB00860.1; -;  
DR InterPro: IPR001225; FA\_desaturase.  
DR Pfam: PF00487; FA\_desaturase; 2.  
DR ProDom: PD001081; FA\_desaturase; 1.  
KW Oxidoreductase; Fatty acid biosynthesis; Endoplasmic reticulum;  
KW Transmembrane.  
FT TRANSMEM 61 81 POTENTIAL.  
FT TRANSMEM 85 105 POTENTIAL.  
FT TRANSMEM 117 137 POTENTIAL.  
FT TRANSMEM 179 199 POTENTIAL.  
FT TRANSMEM 225 245 POTENTIAL.  
FT TRANSMEM 249 269 POTENTIAL.  
FT DOMAIN 105 109 HISTIDINE BOX 1.  
FT DOMAIN 141 145 HISTIDINE BOX 2.  
FT DOMAIN 315 319 HISTIDINE BOX 3.  
SQ SEQUENCE 383 AA; 43967 MW; F23EF7159B2F9967 CRC64;  
  
Query Match 4.3%; Score 102.5; DB 1; Length 383;  
Best Local Similarity 19.7%; Pred. No. 0.8;  
Matches 66; Conservative 45; Mismatches 129; Indels 95; Gaps 15;  
  
Qy 124 TYFLFKATISIMAFYQVLYGWYTSACLLALAWQFGWLTHEFCHQOQTKNRLND 183  
Db 71 THYFLLPGPLSFRGMAYI-----WAV-QGCILTVGVV-----IAHECGHAFESDYQLDD 120  
Qy 184 TISLFFGNFL--QGFSDRWKDKHNTHTAATNVIDHDGIDLAPLFAFIPGD-----LC 235  
Db 121 IVGLILHSALLVPYFS---WKYSRRHSHSNTGSLERDE-----VFVFKOKSCKWYS 169  
Qy 236 KYKASFEKAILKIVPYQHLVFTAML-----PMLRFSWTSQSVQWFKENOMEXKVYQR 288  
Db 170 KYLNPPGPRVLTAVTLTLGWFLYLALNVSGRPYDFACHYDPGYFSIDRE-RLQIYIS 228  
Qy 289 NAFWEQATVGHWAIFYQLFLP-----TWPLRVAYFIISOMGGGLLIAHVVTNHNVS 343  
Db 229 DA-----GVLA-VVYGLFRMAKGLAWVGVVPLLVNGFLVLITLQLHTHPAL 279  
Qy 344 DKYPANSRLNFAALQILITRNMTSPFDIWLWGL-----NYQIE 385  
Db 280 PHYTSSE-----WDWLRLGALATVDRDYGLNLKVFHNITDTHVA 317





**THIS PAGE BLANK (USPTO)**